(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
17 April 2003 (17.04.2003)

PCT

(10) International Publication Number
WO 03/031622 A1

(51) International Patent Classification[7]: C12N 15/29, 15/52, 15/53, 15/55, 15/60, 15/61, A01H 5/00

(21) International Application Number: PCT/AU02/01345

(22) International Filing Date: 4 October 2002 (04.10.2002)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
PR 8113 5 October 2001 (05.10.2001) AU

(71) Applicants (for all designated States except US): AGRICULTURE VICTORIA SERVICES PTY LTD [AU/AU]; 475 Mickleham Road, Attwood, Victoria 3049 (AU). AGRESEARCH LIMITED [NZ/NZ]; 5th Floor, Tower Block, Ruakura Reserach Centre, East Street, Hamilton (NZ).

(72) Inventors; and
(75) Inventors/Applicants (for US only): SPANGENBERG, German [UY/AU]; 56 Arthur Street, Bundoora, Victoria 3083 (AU). SAWBRIDGE, Timothy, Ivor [GB/AU]; 174 Keele Street, Collingwood, Victoria 3066 (AU). ONG, Eng, Kok [MY/AU]; 19 Overland Drive, Vermont South, Victoria 3133 (AU). EMMERLING, Michael [DE/AU]; 249 Nepean Street, Greensborough, Victoria 3088 (AU).

(74) Agent: FREEHILLS CARTER SMITH BEADLE; Level 43, 101 Collins Street, Melbourne, Victoria 3000 (AU).

(81) Designated States (national): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, OM, PH, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.

(84) Designated States (regional): ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, SK, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:
— *with international search report*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

WO 03/031622 A1

(54) Title: MANIPULATION OF FLAVONOID BIOSYNTHESIS IN PLANTS

(57) Abstract: The present invention relates to nucleic acids and nucleic acid fragments encoding amino acid sequences for flavonoid biosynthetic enzymes in plants, and the use thereof for the modification of flavonoid biosynthesis in plants.More particularly, the flavonoid biosynthetic enzyme is selected from the group consisting of chalcone isomerase (CHI), chalcone synthase (CHS), chalcone reductase (CHR), dihydroflavonol 4-reductase (DFR), leucoanthocyanidin reductase (LCR), flavonoid 3', 5' hydrolase (F3'5'H), flavanone 3-hydroxylase (F3H), flavonoid 3'-hydroxylase (F3'H), phenylalanine ammonia-olyase (PAL) and vestitone reductase (VR), and functionally active fragments and variants thereof.

MANIPULATION OF FLAVONOID BIOSYNTHESIS IN PLANTS

The present invention relates to nucleic acids and nucleic acid fragments encoding amino acid sequences for flavonoid biosynthetic enzymes in plants, and the use thereof for the modification of flavonoid biosynthesis in plants.

Flavonoids constitute a relatively diverse family of aromatic molecules that are derived from phenylalanine and malonyl-coenzyme A (CoA, via the fatty acid pathway). These compounds include six major subgroups that are found in most higher plants: the chalcones, flavones, flavonols, flavandiols, anthocyanins and condensed tannins (or proanthocyanidins). A seventh group, the aurones, is widespread, but not ubiquitous.

Some plant species also synthesize specialised forms of flavonoids, such as the isoflavonoids that are found in legumes and a small number of non-legume plants. Similarly, sorghum, maize and gloxinia are among the few species known to synthesize 3-deoxyanthocyanins (or phlobaphenes in the polymerised form). The stilbenes which are closely related to flavonoids, are synthesised by another group of unrelated species that includes grape, peanut and pine.

Besides providing pigmentation to flowers, fruits, seeds, and leaves, flavonoids also have key roles in signalling between plants and microbes, in male fertility of some species, in defense as antimicrobial agents and feeding deterrents, and in UV protection.

Flavonoids also have significant activities when ingested by animals, and there is great interest in their potential health benefits, particularly for compounds such as isoflavonoids, which have been linked to anticancer benefits, and stilbenes that are believed to contribute to reduced heart disease.

The major branch pathways of flavonoid biosynthesis start with general phenylpropanoid metabolism and lead to the nine major subgroups: the colorless chalcones, aurones, isoflavonoids, flavones, flavonols, flavandiols, anthocyanins, condensed tannins, and phlobaphene pigments. The enzyme phenylalanine

ammonia-lyase (PAL) of the general phenylpropanoid pathway will lead to the production of cinnamic acid. Cinnamate-4-hydroxylase (C4H) will produce p-coumaric acid which will be converted through the action of 4-coumaroyl:CoA-ligase (4CL) to the production of 4-coumaroyl-CoA and malonyl-CoA. The first committed step in flavonoid biosynthesis is catalyzed by chalcone synthase (CHS), which uses malonyl CoA and 4-coumaryl CoA as substrates. Chalcone reductase (CHR) balances the production of 5-hydroxy- or 5-deoxyflavonoids. The next enzyme, chalcone isomerase (CHI) catalyses ring closure to form a flavanone, but the reaction can also occur spontaneously. Other enzymes in the pathway are: flavanone 3-hydroxylase (F3H), dihydroflavonol 4-reductase (DFR), flavonoid 3'-hydroxylase (F3'H) and flavonoid 3', 5' hydroxylase (F3'5'H).

The *Arabidopsis BANYULS* gene encodes a DFR-like protein that may be a leucoanthocyanidin reductase (LCR) that catalyzes an early step in condensed tannin biosynthesis. Condensed tannins are plant polyphenols with protein-precipitating and antioxidant properties, synthesized by the flavonoid pathway. Their chemical properties include protein binding, metal chelation, anti-oxidation, and UV-light absorption. As a result condensed tannins inhibit viruses, micro-organisms, insects, fungal pathogens, and monogastric digestion. Moderate amounts of tannins improve forage quality by disrupting protein foam and conferring protection from rumen pasture bloat. Bloat is a digestive disorder that occurs on some highly nutritious forage legumes such as alfalfa (*Medicago sativa*) and white clover (*Trifolium repens*). Moderate amounts of tannin can also reduce digestion rates in the rumen and can reduce parasitic load sufficiently to increase the titre of amino acids and small peptides in the small intestine without compromising total digestion.

Vestitone reductase (VR) is the penultimate enzyme in medicarpin biosynthesis. Medicarpin, a phytoalexin, has been associated with plant resistance to fungal pathogens.

While nucleic acid sequences encoding some flavonoid biosynthetic enzymes CHI, CHS, CHR, DFR, LCR, F3'5'H, F3H, F3'H, PAL and VR have been isolated for certain species of plants, there remains a need for materials useful in

modifying flavonoid biosynthesis; in modifying protein binding, metal chelation, anti-oxidation, and UV-light absorption; in modifying plant pigment production; in modifying plant defense to biotic stresses such as viruses, micro-organisms, insects or fungal pathogens; in modifying forage quality, for example by disrupting protein foam and/or conferring protection from rumen pasture bloat, particularly in forage legumes and grasses, including alfalfa, medics, clovers, ryegrasses and fescues, and for methods for their use.

It is an object of the present invention to overcome, or at least alleviate, one or more of the difficulties or deficiencies associated with the prior art.

In one aspect, the present invention provides substantially purified or isolated nucleic acids or nucleic acid fragments encoding the flavonoid biosynthetic enzymes CHI, CHS, CHR, DFR, LCR, F3'5'H, F3H, F3'H, PAL and VR from a clover (*Trifolium*), medic (*Medicago*), ryegrass (*Lolium*) or fescue (*Festuca*) species and functionally active fragments and variants thereof.

The present invention also provides substantially purified or isolated nucleic acids or nucleic acid fragments encoding amino acid sequences for a class of proteins which are related to CHI, CHS, CHR, DFR, LCR, F3'5'H, F3H, F3'H, PAL and VR and functionally active fragments and variants thereof. Such proteins are referred to herein as CHI-like, CHS-like, CHR-like, DFR-like, LCR-like, F3'5'H-like, F3H-like, F3'H-like, PAL-like and VR-like, respectively.

The individual or simultaneous enhancement or otherwise manipulation of CHI, CHS, CHR, DFR, LCR, F3'5'H, F3H, F3'H, PAL and/or VR or like gene activities in plants may enhance or otherwise alter flavonoid biosynthesis; may enhance or otherwise alter the plant capacity for protein binding, metal chelation, anti-oxidation or UV-light absorption; may enhance or reduce or otherwise alter plant pigment production; may modify plant defense to biotic stresses such as viruses, micro-organisms, insects or fungal pathogens; and/or may modify forage quality, for example by disrupting protein foam and/or conferring protection from rumen pasture bloat.

The individual or simultaneous enhancement or otherwise manipulation of CHI, CHS, CHR, DFR, LCR, F3'5'H, F3H, F3'H, PAL and/or VR or like gene activities in plants has significant consequences for a range of applications in, for example, plant production and plant protection. For example, it has applications in increasing plant tolerance and plant defense to biotic stresses such as viruses, micro-organisms, insects and fungal pathogens; in improving plant forage quality, for example by disrupting protein foam and in conferring protection from rumen pasture bloat; in reducing digestion rates in the rumen and reducing parasitic load; in the production of plant compounds leading to health benefits, such as isoflavonoids, which have been linked to anticancer benefits, and stilbenes that are believed to contribute to reduced heart disease.

Methods for the manipulation of CHI, CHS, CHR, DFR, LCR, F3'5'H, F3H, F3'H, PAL and/or VR or like gene activities in plants, including legumes such as clovers (*Trifolium* species), lucerne (*Medicago sativa*) and grass species such as ryegrasses (*Lolium* species) and fescues (*Festuca* species) may facilitate the production of, for example, forage legumes and forage grasses and other crops with enhanced tolerance to biotic stresses such as viruses, micro-organisms, insects and fungal pathogens; altered pigmentation in flowers; forage legumes with enhanced herbage quality and bloat-safety; crops with enhanced isoflavonoid content leading to health benefits.

The clover (*Trifolium*), medic (*Medicago*), ryegrass (*Lolium*) or fescue (*Festuca*) species may be of any suitable type, including white clover (*Trifolium repens*), red clover (*Trifolium pratense*), subterranean clover (*Trifolium subterraneum*), alfalfa (*Medicago sativa*), Italian or annual ryegrass (*Lolium multiflorum*), perennial ryegrass (*Lolium perenne*), tall fescue (*Festuca arundinacea*), meadow fescue (*Festuca pratensis*) and red fescue (*Festuca rubra*). Preferably the species is a clover or a ryegrass, more preferably white clover (*T. repens*) or perennial ryegrass (*L. perenne*). White clover (*Trifolium repens* L.) and perennial ryegrass (*Lolium perenne* L.) are key pasture legumes and grasses, respectively, in temperate climates throughout the world. Perennial ryegrass is also an important turf grass.

The nucleic acid or nucleic acid fragment may be of any suitable type and includes DNA (such as cDNA or genomic DNA) and RNA (such as mRNA) that is single- or double-stranded, optionally containing synthetic, non-natural or altered nucleotide bases, and combinations thereof.

The term "isolated" means that the material is removed from its original environment (eg. the natural environment if it is naturally occurring). For example, a naturally occurring nucleic acid present in a living plant is not isolated, but the same nucleic acid separated from some or all of the coexisting materials in the natural system, is isolated. Such nucleic acids could be part of a vector and/or such nucleic acids could be part of a composition, and still be isolated in that such a vector or composition is not part of its natural environment.

Such nucleic acids or nucleic acid fragments could be assembled to form a consensus contig. As used herein, the term "consensus contig" refers to a nucleotide sequence that is assembled from two or more constituent nucleotide sequences that share common or overlapping regions of sequence homology. For example, the nucleotide sequence of two or more nucleic acids or nucleic acid fragments can be compared and aligned in order to identify common or overlapping sequences. Where common or overlapping sequences exist between two or more nucleic acids or nucleic acid fragments, the sequences (and thus their corresponding nucleic acids or nucleic acid fragments) can be assembled into a single contiguous nucleotide sequence.

In a preferred embodiment of this aspect of the invention, the substantially purified or isolated nucleic acid or nucleic acid fragment encoding a CHI or CHI-like protein includes a nucleotide sequence selected from the group consisting of (a) sequences shown in Figures 1, 3, 4, 6, 7, 9, 10, 12, 122 and 127 hereto (Sequence ID Nos: 1, 3 to 7, 8, 10 to 12, 13, 15 and 16, 17, 19 to 22, 307, and 309, respectively); (b) complements of the sequences recited in (a); (c) sequences antisense to the sequences recited in (a) and (b); and (d) functionally active fragments and variants of the sequences recited in (a), (b) and (c).

In a further preferred embodiment of this aspect of the invention, the substantially purified or isolated nucleic acid or nucleic acid fragment encoding a CHS or CHS-like protein includes a nucleotide sequence selected from the group consisting of (a) sequences shown in Figures 13, 15, 16, 18, 19, 21, 22, 24, 25, 27, 28, 30, 31, 33, 34, 137, 142, 147, 152, 157 and 162 hereto (Sequence ID Nos: 23, 25 to 63, 64, 66 to 68, 69, 71 to 77, 78, 80 to 90, 91, 93 and 94, 95, 97 to 100, 101, 103 to 105, 106, 313, 315, 317, 319, 321, and 323, respectively); (b) complements of the sequences recited in (a); (c) sequences antisense to the sequences recited in (a) and (b); and (d) functionally active fragments and variants of the sequences recited in (a), (b) and (c).

In a further preferred embodiment of this aspect of the invention, the substantially purified or isolated nucleic acid or nucleic acid fragment encoding a CHR or CHR-like protein includes a nucleotide sequence selected from the group consisting of (a) sequences shown in Figures 36, 38, 40, 41, 43 and 132 hereto (Sequence ID Nos: 108, 110, 112 to 116, 117, 119 to 134, and 311, respectively); (b) complements of the sequences recited in (a); (c) sequences antisense to the sequences recited in (a) and (b); and (d) functionally active fragments and variants of the sequences recited in (a), (b) and (c).

In a further preferred embodiment of this aspect of the invention, the substantially purified or isolated nucleic acid or nucleic acid fragment encoding a DFR or DFR-like protein includes a nucleotide sequence selected from the group consisting of (a) sequences shown in Figures 44, 46, 47, 49, 50, 52, 54, 55, 57, 59, 61, 62, 64, 101, 103, 104, 106, 117 and 167 hereto (Sequence ID Nos: 135, 137 to 146, 147, 149 to 152, 153, 155, 157 and 158, 159, 161, 163, 165 to 167, 168, 170 to 184, 286, 288 to 292, 293, 295 to 297, 305, and 325, respectively); (b) complements of the sequences recited in (a); (c) sequences antisense to the sequences recited in (a) and (b); and (d) functionally active fragments and variants of the sequences recited in (a), (b) and (c).

In a further preferred embodiment of this aspect of the invention, the substantially purified or isolated nucleic acid or nucleic acid fragment encoding an LCR or LCR-like protein includes a nucleotide sequence selected from the group

consisting of (a) sequences shown in Figures 65 and 67 hereto (Sequence ID Nos: 185 and 187 to 193, respectively); (b) complements of the sequences recited in (a); (c) sequences antisense to the sequences recited in (a) and (b); and (d) functionally active fragments and variants of the sequences recited in (a), (b) and (c).

In a still further preferred embodiment of this aspect of the invention, the substantially purified or isolated nucleic acid or nucleic acid fragment encoding an F3'5'H or F3'5'H-like protein includes a nucleotide sequence selected from the group consisting of (a) sequences shown in Figures 68, 70 and 72 hereto (Sequence ID Nos: 194, 196, and 198 to 201, respectively); (b) complements of the sequences recited in (a); (c) sequences antisense to the sequences recited in (a) and (b); and (d) functionally active fragments and variants of the sequences recited in (a), (b) and (c).

In a further preferred embodiment of this aspect of the invention, the substantially purified or isolated nucleic acid or nucleic acid fragment encoding an F3H or F3H-like protein includes a nucleotide sequence selected from the group consisting of (a) sequences shown in Figures 73, 75, 76, 78, 107, 109, 111 and 172 hereto (Sequence ID Nos: 202, 204 to 244, 245, 247, 298, 300 to 302, 303, and 327, respectively); (b) complements of the sequences recited in (a); (c) sequences antisense to the sequences recited in (a) and (b); and (d) functionally active fragments and variants of the sequences recited in (a), (b) and (c).

In a still further preferred embodiment of this aspect of the invention, the substantially purified or isolated nucleic acid or nucleic acid fragment encoding an F3'H or F3'H-like protein includes a nucleotide sequence selected from the group consisting of (a) sequences shown in Figures 80 and 82 hereto (Sequence ID Nos: 249, and 251 and 252, respectively); (b) complements of the sequences recited in (a); (c) sequences antisense to the sequences recited in (a) and (b); and (d) functionally active fragments and variants of the sequences recited in (a), (b) and (c).

In a further preferred embodiment of this aspect of the invention, the substantially purified or isolated nucleic acid or nucleic acid fragment encoding an PAL or PAL-like protein includes a nucleotide sequence selected from the group consisting of (a) sequences shown in Figures 83, 85, 86, 88, 89, 91, 93, 95, 97, 177, 182 and 187 hereto (Sequence ID Nos: 253, 255 to 257, 258, 260 to 267, 268, 270, 272, 274, 276 and 277, 329, 331, and 333, respectively); (b) complements of the sequences recited in (a); (c) sequences antisense to the sequences recited in (a) and (b); and (d) functionally active fragments and variants of the sequences recited in (a), (b) and (c).

In a still further preferred embodiment of this aspect of the invention, the substantially purified or isolated nucleic acid or nucleic acid fragment encoding an VR or VR-like protein includes a nucleotide sequence selected from the group consisting of (a) sequences shown in Figures 98, 100 and 192 hereto (Sequence ID Nos: 278, 280 to 285, and 335, respectively); (b) complements of the sequences recited in (a); (c) sequences antisense to the sequences recited in (a) and (b); and (d) functionally active fragments and variants of the sequences recited in (a), (b) and (c).

By "functionally active" in relation to nucleic acids it is meant that the fragment or variant (such as an analogue, derivative or mutant) encodes a polypeptide which is capable of modifying flavonoid biosynthesis in a plant. Such variants include naturally occurring allelic variants and non-naturally occurring variants. Additions, deletions, substitutions and derivatizations of one or more of the nucleotides are contemplated so long as the modifications do not result in loss of functional activity of the fragment or variant. Preferably the functionally active fragment or variant has at least approximately 80% identity to the relevant part of the above mentioned nucleotide sequence, more preferably at least approximately 90% identity, even more preferably at least approximately 95% identity, most preferably at least approximately 98% homology. Such functionally active variants and fragments include, for example, those having nucleic acid changes which result in conservative amino acid substitutions of one or more residues in the corresponding amino acid sequence. Preferably the fragment has a size of at least

10 nucleotides, more preferably at least 15 nucleotides, most preferably at least 20 nucleotides.

Nucleic acids or nucleic acid fragments encoding at least a portion of several CHI, CHS, CHR, DFR, LCR, F3'5'H, F3H, F3'H, PAL and VR have been isolated and identified. The nucleic acids or nucleic acid fragments of the present invention may be used to isolate cDNAs and genes encoding homologous proteins from the same or other plant species. Isolation of homologous genes using sequence-dependent protocols, such as methods of nucleic acid hybridisation, and methods of DNA and RNA amplification as exemplified by various uses of nucleic acid amplification technologies (e.g. polymerase chain reaction, ligase chain reaction), is well known in the art.

For example, genes encoding other CHI or CHI-like, CHS or CHS-like, CHR or CHR-like, DFR or DFR-like, LCR or LCR-like, F3'5'H or F3'5'H-like, F3H or F3H-like, F3'H or F3'H-like, PAL or PAL-like and VR or VR-like proteins, either as cDNAs or genomic DNAs, may be isolated directly by using all or a portion of the nucleic acids or nucleic acid fragments of the present invention as hybridisation probes to screen libraries from the desired plant employing the methodology well known to those skilled in the art. Specific oligonucleotide probes based upon the nucleic acid sequences of the present invention may be designed and synthesized by methods known in the art. Moreover, the entire sequences may be used directly to synthesize DNA probes by methods known to the skilled artisan such as random primer DNA labelling, nick translation, or end-labelling techniques, or RNA probes using available *in vitro* transcription systems. In addition, specific primers may be designed and used to amplify a part or all of the sequences of the present invention. The resulting amplification products may be labelled directly during amplification reactions or labelled after amplification reactions, and used as probes to isolate full-length cDNA or genomic fragments under conditions of appropriate stringency.

In addition, short segments of the nucleic acids or nucleic acid fragments of the present invention may be used in amplification protocols to amplify longer nucleic acids or nucleic acid fragments encoding homologous genes from DNA or

RNA. For example, polymerase chain reaction may be performed on a library of cloned nucleic acid fragments wherein the sequence of one primer is derived from the nucleic acid sequences of the present invention, and the sequence of the other primer takes advantage of the presence of the polyadenylic acid tracts to the 3' end of the mRNA precursor encoding plant genes. Alternatively, the second primer sequence may be based upon sequences derived from the cloning vector. For example, those skilled in the art can follow the RACE protocol [Frohman *et al.* (1988) *Proc. Natl. Acad Sci. USA* 85:8998, the entire disclosure of which is incorporated herein by reference] to generate cDNAs by using PCR to amplify copies of the region between a single point in the transcript and the 3' or 5' end. Using commercially available 3' RACE and 5' RACE systems (BRL), specific 3' or 5' cDNA fragments may be isolated [Ohara *et al.* (1989) *Proc. Natl. Acad Sci USA* 86:5673; Loh *et al.* (1989) *Science* 243:217, the entire disclosures of which are incorporated herein by reference]. Products generated by the 3' and 5' RACE procedures may be combined to generate full-length cDNAs.

In a second aspect of the present invention there is provided a substantially purified or isolated polypeptide from a clover (*Trifolium*), medic (*Medicago*), ryegrass (*Lolium*) or fescue (*Festuca*) species, selected from the group consisting of CHI and CHI-like, CHS and CHS-like, CHR and CHR-like, DFR and DFR-like, LCR and LCR-like, F3'5'H and F3'5'H-like, F3H and F3H-like, F3'H and F3'H-like, PAL and PAL-like, VR and VR-like; and functionally active fragments and variants thereof.

The clover (*Trifolium*), medic (*Medicago*), ryegrass (*Lolium*) or fescue (*Festuca*) species may be of any suitable type, including white clover (*Trifolium repens*), red clover (*Trifolium pratense*), subterranean clover (*Trifolium subterraneum*), alfalfa (*Medicago sativa*), Italian or annual ryegrass (*Lolium multiflorum*), perennial ryegrass (*Lolium perenne*), tall fescue (*Festuca arundinacea*), meadow fescue (*Festuca pratensis*) and red fescue (*Festuca rubra*). Preferably the species is a clover or a ryegrass, more preferably white clover (*T. repens*) or perennial ryegrass (*L. perenne*).

In a preferred embodiment of this aspect of the invention, the substantially purified or isolated CHI or CHI-like polypeptide includes an amino acid sequence selected from the group consisting of the sequences shown in Figures 2, 5, 8, 11, 123 and 128 hereto (Sequence ID Nos: 2, 9, 14, 18, 308, and 310, respectively), and functionally active fragments and variants thereof.

In a further preferred embodiment of this aspect of the invention, the substantially purified or isolated CHS or CHS-like polypeptide includes an amino acid sequence selected from the group consisting of the sequences shown in Figures 14, 17, 20, 23, 26, 29, 32, 35, 138, 143, 148, 153, 158 and 163 hereto (Sequence ID Nos: 24, 65, 70, 79, 92, 96, 102, 107, 314, 316, 318, 320, 322, and 324, respectively), and functionally active fragments and variants thereof.

In a further preferred embodiment of this aspect of the invention, the substantially purified or isolated CHR or CHR-like polypeptide includes an amino acid sequence selected from the group consisting of the sequences shown in Figures 37, 39, 42 and 133 hereto (Sequence ID Nos: 109, 111, 118, and 312, respectively), and functionally active fragments and variants thereof.

In a still further preferred embodiment of this aspect of the invention, the substantially purified or isolated DFR or DFR-like polypeptide includes an amino acid sequence selected from the group consisting of the sequences shown in Figures 45, 48, 51, 53, 56, 58, 60, 63, 102, 105, 118 and 168 hereto (Sequence ID Nos: 136, 148, 54, 156, 160, 162, 164, 169, 287, 294, 306, and 326, respectively), and functionally active fragments and variants thereof.

In a still further preferred embodiment of this aspect of the invention, the substantially purified or isolated LCR or LCR-like polypeptide includes an amino acid sequence shown in Figure 66 hereto (Sequence ID No: 186), and functionally active fragments and variants thereof.

In a still further preferred embodiment of this aspect of the invention, the substantially purified or isolated F3'5'H or F3'5'H-like polypeptide includes an amino acid sequence selected from the group consisting of the sequences shown

in Figures 69 and 71 hereto (Sequence ID Nos: 195 and 197, respectively), and functionally active fragments and variants thereof.

In a still further preferred embodiment of this aspect of the invention, the substantially purified or isolated F3H or F3H-like polypeptide includes an amino acid sequence selected from the group consisting of the sequences shown in Figures 74, 77, 79, 108, 112 and 173 hereto (Sequence ID Nos: 203, 246, 248, 299, 304, and 328, respectively), and functionally active fragments and variants thereof.

In a still further preferred embodiment of this aspect of the invention, the substantially purified or isolated F3'H or F3'H-like polypeptide includes an amino acid sequence shown in Figure 81 hereto (Sequence ID No: 250), and functionally active fragments and variants thereof.

In a still further preferred embodiment of this aspect of the invention, the substantially purified or isolated PAL or PAL-like polypeptide includes an amino acid sequence selected from the group consisting of the sequences shown in Figures 84, 87, 90, 92, 94, 96, 178, 183 and 188 hereto (Sequence ID Nos: 254, 259, 269, 271, 273, 275, 330, 332, and 334, respectively), and functionally active fragments and variants thereof.

In a still further preferred embodiment of this aspect of the invention, the substantially purified or isolated VR or VR-like polypeptide includes an amino acid sequence shown in Figures 99 and 193 hereto (Sequence ID Nos: 279 and 336, respectively), and functionally active fragments and variants thereof.

By "functionally active" in relation to polypeptides it is meant that the fragment or variant has one or more of the biological properties of the proteins CHI, CHI-like, CHS, CHS-like, CHR, CHR-like, DFR, DFR-like, LCR, LCR-like, F3'5'H, F3'5'H-like, F3H, F3H-like, F3'H, F3'H-like, PAL, PAL-like, VR and VR-like, respectively. Additions, deletions, substitutions and derivatizations of one or more of the amino acids are contemplated so long as the modifications do not result in loss of functional activity of the fragment or variant. Preferably the functionally

active fragment or variant has at least approximately 60% identity to the relevant part of the above mentioned amino acid sequence, more preferably at least approximately 80% identity, even more preferably at least approximately 90% identity most preferably at least approximately 95% homology. Such functionally active variants and fragments include, for example, those having conservative amino acid substitutions of one or more residues in the corresponding amino acid sequence. Preferably the fragment has a size of at least 10 amino acids, more preferably at least 15 amino acids, most preferably at least 20 amino acids.

In a further embodiment of this aspect of the invention, there is provided a polypeptide recombinantly produced from a nucleic acid or nucleic acid fragment according to the present invention. Techniques for recombinantly producing polypeptides are well known to those skilled in the art.

Availability of the nucleotide sequences of the present invention and deduced amino acid sequences facilitates immunological screening of cDNA expression libraries. Synthetic peptides representing portions of the instant amino acid sequences may be synthesized. These peptides may be used to immunise animals to produce polyclonal or monoclonal antibodies with specificity for peptides and/or proteins including the amino acid sequences. These antibodies may be then used to screen cDNA expression libraries to isolate full-length cDNA clones of interest.

A genotype is the genetic constitution of an individual or group. Variations in genotype are important in commercial breeding programs, in determining parentage, in diagnostics and fingerprinting, and the like. Genotypes can be readily described in terms of genetic markers. A genetic marker identifies a specific region or locus in the genome. The more genetic markers, the finer defined is the genotype. A genetic marker becomes particularly useful when it is allelic between organisms because it then may serve to unambiguously identify an individual. Furthermore, a genetic marker becomes particularly useful when it is based on nucleic acid sequence information that can unambiguously establish a genotype of an individual and when the function encoded by such nucleic acid is known and is associated with a specific trait. Such nucleic acids and/or nucleotide

sequence information including single nucleotide polymorphisms (SNPs), variations in single nucleotides between allelic forms of such nucleotide sequence, may be used as perfect markers or candidate genes for the given trait.

Applicants have identified a number of SNPs of the nucleic acids or nucleic acid fragments of the present invention. These are indicated (marked with grey on the black background) in the figures that show multiple alignments of nucleotide sequences of nucleic acid fragments contributing to consensus contig sequences. See for example, Figures 3, 6, 9, 12, 15, 18, 21, 24, 27, 30, 33, 40, 43, 46, 49, 54, 61, 64, 67, 72, 75, 82, 85, 88, 97, 100, 103, 106 and 109 hereto (Sequence ID Nos: 3 to 7, 10 to 12, 15 and 16, 19 to 22, 25 to 63, 66 to 68, 71 to 77, 80 to 90, 93 and 94, 97 to 100, 103 to 105, 112 to 116, 119 to 134, 137 to 146, 149 to 152, 157 and 158, 165 to 167, 170 to 184, 187 to 193, 198 to 201, 204 to 244, 251 and 252, 255 to 257, 260 to 267, 276 and 277, 280 to 285, 288 to 292, 295 to 297, and 300 to 302, respectively).

Accordingly, in a further aspect of the present invention, there is provided a substantially purified or isolated nucleic acid or nucleic acid fragment including a single nucleotide polymorphism (SNP) from a nucleic acid or nucleic acid fragment according to the present invention or complements or sequences antisense thereto, and functionally active fragments and variants thereof.

In a still further aspect of the present invention there is provided a method of isolating a nucleic acid or nucleic acid fragment of the present invention including a SNP, said method including sequencing nucleic acid fragments from a nucleic acid library.

The nucleic acid library may be of any suitable type and is preferably a cDNA library.

The nucleic acid or nucleic acid fragments may be isolated from a recombinant plasmid or may be amplified, for example using polymerase chain reaction.

The sequencing may be performed by techniques known to those skilled in the art.

In a still further aspect of the present invention, there is provided use of the nucleic acids or nucleic acid fragments of the present invention including SNPs, and/or nucleotide sequence information thereof, as molecular genetic markers.

In a still further aspect of the present invention there is provided use of a nucleic acid or nucleic acid fragment of the present invention, and/or nucleotide sequence information thereof, as a molecular genetic marker.

More particularly, nucleic acids or nucleic acid fragments according to the present invention and/or nucleotide sequence information thereof may be used as a molecular genetic marker for quantitative trait loci (QTL) tagging, QTL mapping, DNA fingerprinting and in marker assisted selection, particularly in clovers, alfalfa, ryegrasses and fescues. Even more particularly, nucleic acids or nucleic acid fragments according to the present invention and/or nucleotide sequence information thereof may be used as molecular genetic markers in plant improvement in relation to plant tolerance to biotic stresses such as viruses, micro-organisms, insects, fungal pathogens; in relation to forage quality; in relation to bloat safety; in relation to condensed tannin content; in relation to plant pigmentation. Even more particularly, sequence information revealing SNPs in allelic variants of the nucleic acids or nucleic acid fragments of the present invention and/or nucleotide sequence information thereof may be used as molecular genetic markers for QTL tagging and mapping and in marker assisted selection, particularly in clovers, alfalfa, ryegrasses and fescues.

In a still further aspect of the present invention there is provided a construct including a nucleic acid or nucleic acid fragment according to the present invention.

The term "construct" as used herein refers to an artificially assembled or isolated nucleic acid molecule which includes the gene of interest. In general a construct may include the gene or genes of interest, a marker gene which in some

cases can also be the gene of interest and appropriate regulatory sequences. It should be appreciated that the inclusion of regulatory sequences in a construct is optional, for example, such sequences may not be required in situations where the regulatory sequences of a host cell are to be used. The term construct includes vectors but should not be seen as being limited thereto.

In a still further aspect of the present invention there is provided a vector including a nucleic acid or nucleic acid fragment according to the present invention.

The term "vector" as used herein encompasses both cloning and expression vectors. Vectors are often recombinant molecules containing nucleic acid molecules from several sources.

In a preferred embodiment of this aspect of the invention, the vector may include a regulatory element such as a promoter, a nucleic acid or nucleic acid fragment according to the present invention and a terminator; said regulatory element, nucleic acid or nucleic acid fragment and terminator being operatively linked.

By "operatively linked" is meant that said regulatory element is capable of causing expression of said nucleic acid or nucleic acid fragment in a plant cell and said terminator is capable of terminating expression of said nucleic acid or nucleic acid fragment in a plant cell. Preferably, said regulatory element is upstream of said nucleic acid or nucleic acid fragment and said terminator is downstream of said nucleic acid or nucleic acid fragment.

The vector may be of any suitable type and may be viral or non-viral. The vector may be an expression vector. Such vectors include chromosomal, non-chromosomal and synthetic nucleic acid sequences, eg. derivatives of plant viruses; bacterial plasmids; derivatives of the Ti plasmid from *Agrobacterium tumefaciens*, derivatives of the Ri plasmid from *Agrobacterium rhizogenes*; phage DNA; yeast artificial chromosomes; bacterial artificial chromosomes; binary bacterial artificial chromosomes; vectors derived from combinations of plasmids

and phage DNA. However, any other vector may be used as long as it is replicable, integrative or viable in the plant cell.

The regulatory element and terminator may be of any suitable type and may be endogenous to the target plant cell or may be exogenous, provided that they are functional in the target plant cell.

Preferably the regulatory element is a promoter. A variety of promoters which may be employed in the vectors of the present invention are well known to those skilled in the art. Factors influencing the choice of promoter include the desired tissue specificity of the vector, and whether constitutive or inducible expression is desired and the nature of the plant cell to be transformed (eg. monocotyledon or dicotyledon). Particularly suitable constitutive promoters include the Cauliflower Mosaic Virus 35S (CaMV 35S) promoter and derivatives thereof, the maize Ubiquitin promoter, and the rice Actin promoter.

A variety of terminators which may be employed in the vectors of the present invention are also well known to those skilled in the art. The terminator may be from the same gene as the promoter sequence or a different gene. Particularly suitable terminators are polyadenylation signals, such as the CaMV 35S polyA and other terminators from the nopaline synthase (*nos*), the octopine synthase (*ocs*) and the rbcS genes.

The vector, in addition to the regulatory element, the nucleic acid or nucleic acid fragment of the present invention and the terminator, may include further elements necessary for expression of the nucleic acid or nucleic acid fragment, in different combinations, for example vector backbone, origin of replication (ori), multiple cloning sites, spacer sequences, enhancers, introns (such as the maize Ubiquitin *Ubi* intron), antibiotic resistance genes and other selectable marker genes [such as the neomycin phosphotransferase (*npt2*) gene, the hygromycin phosphotransferase (*hph*) gene, the phosphinothricin acetyltransferase (*bar* or *pat*) gene and the gentamycin acetyl transferase (*aacC1*) gene], and reporter genes [such as beta-glucuronidase (GUS) gene (*gusA*) and green fluorescent protein

(gfp)]. The vector may also contain a ribosome binding site for translation initiation. The vector may also include appropriate sequences for amplifying expression.

As an alternative to use of a selectable marker gene to provide a phenotypic trait for selection of transformed host cells, the presence of the vector in transformed cells may be determined by other techniques well known in the art, such as PCR (polymerase chain reaction), Southern blot hybridisation analysis, histochemical GUS assays, northern and Western blot hybridisation analyses.

Those skilled in the art will appreciate that the various components of the vector are operatively linked, so as to result in expression of said nucleic acid or nucleic acid fragment. Techniques for operatively linking the components of the vector of the present invention are well known to those skilled in the art. Such techniques include the use of linkers, such as synthetic linkers, for example including one or more restriction enzyme sites.

The constructs and vectors of the present invention may be incorporated into a variety of plants, including monocotyledons (such as grasses from the genera *Lolium*, *Festuca*, *Paspalum*, *Pennisetum*, *Panicum* and other forage and turfgrasses, corn, oat, sugarcane, wheat and barley), dicotyledons (such as Arabidopsis, tobacco, clovers, medics, eucalyptus, potato, sugarbeet, canola, soybean, chickpea) and gymnosperms. In a preferred embodiment, the constructs and vectors may be used to transform monocotyledons, preferably grass species such as ryegrasses (*Lolium* species) and fescues (*Festuca* species), more preferably perennial ryegrass, including forage- and turf-type cultivars. In an alternate preferred embodiment, the constructs and vectors may be used to transform dicotyledons, preferably forage legume species such as clovers (*Trifolium* species) and medics (*Medicago* species), more preferably white clover (*Trifolium repens*), red clover (*Trifolium pratense*), subterranean clover (*Trifolium subterraneum*) and alfalfa (*Medicago sativa*). Clovers, alfalfa and medics are key pasture legumes in temperate climates throughout the world.

Techniques for incorporating the constructs and vectors of the present invention into plant cells (for example by transduction, transfection or

transformation) are well known to those skilled in the art. Such techniques include *Agrobacterium* mediated introduction, electroporation to tissues, cells and protoplasts, protoplast fusion, injection into reproductive organs, injection into immature embryos and high velocity projectile introduction to cells, tissues, calli, immature and mature embryos. The choice of technique will depend largely on the type of plant to be transformed.

Cells incorporating the constructs and vectors of the present invention may be selected, as described above, and then cultured in an appropriate medium to regenerate transformed plants, using techniques well known in the art. The culture conditions, such as temperature, pH and the like, will be apparent to the person skilled in the art. The resulting plants may be reproduced, either sexually or asexually, using methods well known in the art, to produce successive generations of transformed plants.

In a further aspect of the present invention there is provided a plant cell, plant, plant seed or other plant part, including, e.g. transformed with, a construct, vector, nucleic acid or nucleic acid fragment of the present invention.

The plant cell, plant, plant seed or other plant part may be from any suitable species, including monocotyledons, dicotyledons and gymnosperms. In a preferred embodiment the plant cell, plant, plant seed or other plant part may be from a monocotyledon, preferably a grass species, more preferably a ryegrass (*Lolium* species) or fescue (*Festuca* species), even more preferably perennial ryegrass, including both forage- and turf-type cultivars. In an alternate preferred embodiment the plant cell, plant, plant seed or other plant part may be from a dicotyledon, preferably forage legume species such as clovers (*Trifolium* species) and medics (*Medicago* species), more preferably white clover (*Trifolium repens*), red clover (*Trifolium pratense*), subterranean clover (*Trifolium subterraneum*) and alfalfa (*Medicago sativa*).

The present invention also provides a plant, plant seed or other plant part, or a plant extract derived from a plant cell of the present invention.

The present invention also provides a plant, plant seed or other plant part, or a plant extract derived from a plant of the present invention.

In a further aspect of the present invention there is provided a method of modifying flavonoid biosynthesis in a plant; said method including introducing into said plant an effective amount of a nucleic acid or nucleic acid fragment, construct and/or a vector according to the present invention.

In a further aspect of the present invention there is provided a method of modifying protein binding, metal chelation, anti-oxidation, and/or UV-light absorption in a plant, said method including introducing into said plant an effective amount of a nucleic acid or nucleic acid fragment, construct and/or a vector according to the present invention.

In a further aspect of the present invention there is provided a method of modifying pigment production in a plant, said method including introducing into said plant an effective amount of a nucleic acid or nucleic acid fragment, construct and/or a vector according to the present invention.

In a further aspect of the present invention there is provided a method of modifying plant defense to biotic stresses such as viruses, micro-organisms, insects and fungal pathogens, said method including introducing into said plant an effective amount of a nucleic acid or nucleic acid fragment, construct and/or a vector according to the present invention.

In a further aspect of the present invention there is provided a method of modifying forage quality of a plant by disrupting protein foam and/or conferring protection from rumen pasture bloat, said method including introducing into said plant an effective amount of a nucleic acid or nucleic acid fragment, construct and/or a vector according to the present invention.

By "an effective amount" it is meant an amount sufficient to result in an identifiable phenotypic trait in said plant, or a plant, plant seed or other plant part derived therefrom. Such amounts can be readily determined by an appropriately

skilled person, taking into account the type of plant, the route of administration and other relevant factors. Such a person will readily be able to determine a suitable amount and method of administration. See, for example, Maniatis et al, Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory, Cold Spring Harbor, the entire disclosure of which is incorporated herein by reference.

Using the methods and materials of the present invention, flavonoid biosynthesis, protein binding, metal chelation, anti-oxidation, UV-light absorption, tolerance to biotic stresses such as viruses, micro-organisms, insects and fungal pathogens; pigmentation in for example flowers and leaves; herbage quality and bloat-safety; and/or isoflavonoid content leading to health benefits, may be increased or otherwise modified, for example by incorporating additional copies of a sense nucleic acid or nucleic acid fragment of the present invention. They may be decreased or otherwise modified, for example by incorporating an antisense nucleic acid or nucleic acid fragment of the present invention.

The present invention will now be more fully described with reference to the accompanying Examples and drawings. It should be understood, however, that the description following is illustrative only and should not be taken in any way as a restriction on the generality of the invention described above.

In the Figures

Figure 1 shows the consensus contig nucleotide sequence of TrCHIa (Sequence ID No: 1).

Figure 2 shows the deduced amino acid sequence of TrCHIa (Sequence ID No: 2).

Figure 3 shows the nucleotide sequences of the nucleic acid fragments contributing to the consensus contig sequence TrCHIa (Sequence ID Nos: 3 to 7).

Figure 4 shows the consensus contig nucleotide sequence of TrCHIb (Sequence ID No: 8).

Figure 5 shows the deduced amino acid sequence of TrCHIb (Sequence ID No: 9).

Figure 6 shows the nucleotide sequences of the nucleic acid fragments contributing to the consensus contig sequence TrCHIb (Sequence ID Nos: 10 to 12).

Figure 7 shows the consensus contig nucleotide sequence of TrCHIc (Sequence ID No: 13).

Figure 8 shows the deduced amino acid sequence of TrCHIc (Sequence ID No: 14).

Figure 9 shows the nucleotide sequences of the nucleic acid fragments contributing to the consensus contig sequence TrCHIc (Sequence ID Nos: 15 and 16).

Figure 10 shows the consensus contig nucleotide sequence of TrCHId (Sequence ID No: 17).

Figure 11 shows the deduced amino acid sequence of TrCHId (Sequence ID No: 18).

Figure 12 shows the nucleotide sequences of the nucleic acid fragments contributing to the consensus contig sequence TrCHId (Sequence ID Nos: 19 to 22).

Figure 13 shows the consensus contig nucleotide sequence of TrCHSa (Sequence ID No: 23).

Figure 14 shows the deduced amino acid sequence of TrCHSa (Sequence ID No: 24).

Figure 15 shows the nucleotide sequences of the nucleic acid fragments contributing to the consensus contig sequence TrCHSa (Sequence ID Nos: 25 to 63).

Figure 16 shows the consensus contig nucleotide sequence of TrCHSb (Sequence ID No: 64).

Figure 17 shows the deduced amino acid sequence of TrCHSb (Sequence ID No: 65).

Figure 18 shows the nucleotide sequences of the nucleic acid fragments contributing to the consensus contig sequence TrCHSb (Sequence ID Nos: 66 to 68).

Figure 19 shows the consensus contig nucleotide sequence of TrCHSc (Sequence ID No: 69).

Figure 20 shows the deduced amino acid sequence of TrCHSc (Sequence ID No: 70).

Figure 21 shows the nucleotide sequences of the nucleic acid fragments contributing to the consensus contig sequence TrCHSc (Sequence ID Nos: 71 to 77).

Figure 22 shows the consensus contig nucleotide sequence of TrCHSd (Sequence ID No: 78).

Figure 23 shows the deduced amino acid sequence of TrCHSd (Sequence ID No: 79).

Figure 24 shows the nucleotide sequences of the nucleic acid fragments contributing to the consensus contig sequence TrCHSd (Sequence ID Nos: 80 to 90).

Figure 25 shows the consensus contig nucleotide sequence of TrCHSe (Sequence ID No: 91).

Figure 26 shows the deduced amino acid sequence of TrCHSe (Sequence ID No: 92).

Figure 27 shows the nucleotide sequences of the nucleic acid fragments contributing to the consensus contig sequence TrCHSe (Sequence ID Nos: 93 and 94).

Figure 28 shows the consensus contig nucleotide sequence of TrCHSf (Sequence ID No: 95).

Figure 29 shows the deduced amino acid sequence of TrCHSf (Sequence ID No: 96).

Figure 30 shows the nucleotide sequences of the nucleic acid fragments contributing to the consensus contig sequence TrCHSf (Sequence ID Nos: 97 to 100).

Figure 31 shows the consensus contig nucleotide sequence of TrCHSg (Sequence ID No: 101).

Figure 32 shows the deduced amino acid sequence of TrCHSg (Sequence ID No: 102).

Figure 33 shows the nucleotide sequences of the nucleic acid fragments contributing to the consensus contig sequence TrCHSg (Sequence ID Nos: 103 to 105).

Figure 34 shows the consensus contig nucleotide sequence of TrCHSh (Sequence ID No: 106).

Figure 35 shows the deduced amino acid sequence of TrCHSh (Sequence ID No: 107).

Figure 36 shows the nucleotide sequence of TrCHRa (Sequence ID No: 108).

Figure 37 shows the deduced amino acid sequence of TrCHRa (Sequence ID No: 109).

Figure 38 shows the consensus contig nucleotide sequence of TrCHRb (Sequence ID No: 110).

Figure 39 shows the deduced amino acid sequence of TrCHRb (Sequence ID No: 111).

Figure 40 shows the nucleotide sequences of the nucleic acid fragments contributing to the consensus contig sequence TrCHRb (Sequence ID Nos: 112 to 116).

Figure 41 shows the consensus contig nucleotide sequence of TrCHRc (Sequence ID No: 117).

Figure 42 shows the deduced amino acid sequence of TrCHRc (Sequence ID No: 118).

Figure 43 shows the nucleotide sequences of the nucleic acid fragments contributing to the consensus contig sequence TrCHRc (Sequence ID Nos: 119 to 134).

Figure 44 shows the consensus contig nucleotide sequence of TrDFRa (Sequence ID No: 135).

Figure 45 shows the deduced amino acid sequence of TrDFRa (Sequence ID No: 136).

Figure 46 shows the nucleotide sequences of the nucleic acid fragments contributing to the consensus contig sequence TrDFRa (Sequence ID Nos: 137 to 146).

Figure 47 shows the consensus contig nucleotide sequence of TrDFRb (Sequence ID No: 147).

Figure 48 shows the deduced amino acid sequence of TrDFRb (Sequence ID No: 148).

Figure 49 shows the nucleotide sequences of the nucleic acid fragments contributing to the consensus contig sequence TrDFRb (Sequence ID Nos: 149 to 152).

Figure 50 shows the nucleotide sequence of TrDFRc (Sequence ID No: 153).

Figure 51 shows the deduced amino acid sequence of TrDFRc (Sequence ID No: 154).

Figure 52 shows the consensus contig nucleotide sequence of TrDFRd (Sequence ID No: 155).

Figure 53 shows the deduced amino acid sequence of TrDFRd (Sequence ID No: 156).

Figure 54 shows the nucleotide sequences of the nucleic acid fragments contributing to the consensus contig sequence TrDFRd (Sequence ID Nos: 157 and 158).

Figure 55 shows the nucleotide sequence of TrDFRe (Sequence ID No: 159).

Figure 56 shows the deduced amino acid sequence of TrDFRe (Sequence ID No: 160).

Figure 57 shows the nucleotide sequence of TrDFRf (Sequence ID No: 161).

Figure 58 shows the deduced amino acid sequence of TrDFRf (Sequence ID No: 162).

Figure 59 shows the consensus contig nucleotide sequence of TrDFRg (Sequence ID No: 163).

Figure 60 shows the deduced amino acid sequence of TrDFRg (Sequence ID No: 164).

Figure 61 shows the nucleotide sequences of the nucleic acid fragments contributing to the consensus contig sequence TrDFRg (Sequence ID Nos: 165 to 167).

Figure 62 shows the consensus contig nucleotide sequence of TrDFRh (Sequence ID No: 168).

Figure 63 shows the deduced amino acid sequence of TrDFRh (Sequence ID No: 169).

Figure 64 shows the nucleotide sequences of the nucleic acid fragments contributing to the consensus contig sequence TrDFRh (Sequence ID Nos: 170 to 184).

Figure 65 shows the consensus contig nucleotide sequence of TrLCRa (Sequence ID No: 185).

Figure 66 shows the deduced amino acid sequence of TrLCRa (Sequence ID No: 186).

Figure 67 shows the nucleotide sequences of the nucleic acid fragments contributing to the consensus contig sequence TrLCRa (Sequence ID Nos: 187 to 193).

Figure 68 shows the nucleotide sequence of TrF3'5'Ha (Sequence ID No: 194).

Figure 69 shows the deduced amino acid sequence of TrF3'5'Ha (Sequence ID No: 195).

Figure 70 shows the consensus contig nucleotide sequence of TrF3'5'Hb (Sequence ID No: 196).

Figure 71 shows the deduced amino acid sequence of TrF3'5'Hb (Sequence ID No: 197).

Figure 72 shows the nucleotide sequences of the nucleic acid fragments contributing to the consensus contig sequence TrF3'5'Hb (Sequence ID Nos: 198 to 201).

Figure 73 shows the consensus contig nucleotide sequence of TrF3Ha (Sequence ID No: 202).

Figure 74 shows the deduced amino acid sequence of TrF3Ha (Sequence ID No: 203).

Figure 75 shows the nucleotide sequences of the nucleic acid fragments contributing to the consensus contig sequence TrF3Ha (Sequence ID Nos: 204 to 244).

Figure 76 shows the nucleotide sequence of TrF3Hb (Sequence ID No: 245).

Figure 77 shows the deduced amino acid sequence of TrF3Hb (Sequence ID No: 246).

Figure 78 shows the nucleotide sequence of TrF3Hc (Sequence ID No: 247).

Figure 79 shows the deduced amino acid sequence of TrF3Hc (Sequence ID No: 248).

Figure 80 shows the consensus contig nucleotide sequence of TrF3'Ha (Sequence ID No: 249).

Figure 81 shows the deduced amino acid sequence of TrF3'Ha (Sequence ID No: 250).

Figure 82 shows the nucleotide sequences of the nucleic acid fragments contributing to the consensus contig sequence TrF3'Ha (Sequence ID Nos: 251 and 252).

Figure 83 shows the consensus contig nucleotide sequence of TrPALa (Sequence ID No: 253).

Figure 84 shows the deduced amino acid sequence of TrPALa (Sequence ID No: 254).

Figure 85 shows the nucleotide sequences of the nucleic acid fragments contributing to the consensus contig sequence TrPALa (Sequence ID Nos: 255 to 257).

Figure 86 shows the consensus contig nucleotide sequence of TrPALb (Sequence ID No: 258).

Figure 87 shows the deduced amino acid sequence of TrPALb (Sequence ID No: 259).

Figure 88 shows the nucleotide sequences of the nucleic acid fragments contributing to the consensus contig sequence TrPALb (Sequence ID Nos: 260 to 267).

Figure 89 shows the nucleotide sequence of TrPALc (Sequence ID No: 268).

Figure 90 shows the deduced amino acid sequence of TrPALc (Sequence ID No: 269).

Figure 91 shows the nucleotide sequence of TrPALd (Sequence ID No: 270).

Figure 92 shows the deduced amino acid sequence of TrPALd (Sequence ID No: 271).

Figure 93 shows the nucleotide sequence of TrPALe (Sequence ID No: 272).

Figure 94 shows the deduced amino acid sequence of TrPALe (Sequence ID No: 273).

Figure 95 shows the consensus contig nucleotide sequence of TrPALf (Sequence ID No: 274).

Figure 96 shows the deduced amino acid sequence of TrPALf (Sequence ID No: 275).

Figure 97 shows the nucleotide sequences of the nucleic acid fragments contributing to the consensus contig sequence TrPALf (Sequence ID Nos: 276 and 277).

Figure 98 shows the consensus contig nucleotide sequence of TrVRa (Sequence ID No: 278).

Figure 99 shows the deduced amino acid sequence of TrVRa (Sequence ID No: 279).

Figure 100 shows the nucleotide sequences of the nucleic acid fragments contributing to the consensus contig sequence TrVRa (Sequence ID Nos: 280 to 285).

Figure 101 shows the consensus contig nucleotide sequence of LpDFRa (Sequence ID No: 286).

Figure 102 shows the deduced amino acid sequence of LpDFRa (Sequence ID No: 287).

Figure 103 shows the nucleotide sequences of the nucleic acid fragments contributing to the consensus contig sequence LpDFRa (Sequence ID Nos: 288 to 292).

Figure 104 shows the consensus contig nucleotide sequence of LpDFRb (Sequence ID No: 293).

Figure 105 shows the deduced amino acid sequence of LpDFRb (Sequence ID No: 294).

Figure 106 shows the nucleotide sequences of the nucleic acid fragments contributing to the consensus contig sequence LpDFRb (Sequence ID Nos: 295 to 297).

Figure 107 shows the consensus contig nucleotide sequence of LpF3Ha (Sequence ID No: 298).

Figure 108 shows the deduced amino acid sequence of LpF3Ha (Sequence ID No: 299).

Figure 109 shows the nucleotide sequences of the nucleic acid fragments contributing to the consensus contig sequence LpF3Ha (Sequence ID Nos: 300 to 302).

Figure 110 shows a plasmid map of the cDNA encoding perennial ryegrass F3OH.

Figure 111 shows the full nucleotide sequence of perennial ryegrass F3OH cDNA (Sequence ID No: 303).

Figure 112 shows the deduced amino acid sequence of perennial ryegrass F3OH cDNA (Sequence ID No: 304).

Figure 113 shows plasmid maps of sense and antisense constructs of LpF3OH in pDH51 transformation vector.

Figure 114 shows plasmid maps of sense and antisense constructs of LpF3OH in pPZP221:35S^2 binary transformation vector.

Figure 115 shows screening by Southern hybridisation for RFLPs using LpF3OH as a probe.

Figure 116 shows a plasmid map of the cDNA encoding white clover BANa.

Figure 117 shows the full nucleotide sequence of white clover BANa cDNA (Sequence ID No: 305).

Figure 118 shows the deduced amino acid sequence of white clover BANa cDNA (Sequence ID No: 306).

Figure 119 shows plasmid maps of sense and antisense constructs of TrBANa in pDH51 transformation vector.

Figure 120 shows plasmid maps of sense and antisense constructs of TrBANa in pPZP221:35S^2 binary transformation vector.

Figure 121 shows a plasmid map of the cDNA encoding white clover CHIa.

Figure 122 shows the full nucleotide sequence of white clover CHIa cDNA (Sequence ID No: 307).

Figure 123 shows the deduced amino acid sequence of white clover CHIa cDNA (Sequence ID No: 308).

Figure 124 shows plasmid maps of sense and antisense constructs of TrCHIa in pDH51 transformation vector.

Figure 125 shows plasmid maps of sense and antisense constructs of TrCHIa in pPZP221:35S^2 binary transformation vector.

Figure 126 shows a plasmid map of the cDNA encoding white clover CHId.

Figure 127 shows the full nucleotide sequence of white clover CHId cDNA (Sequence ID No: 309).

Figure 128 shows the deduced amino acid sequence of white clover CHId cDNA (Sequence ID No: 310).

Figure 129 shows plasmid maps of sense and antisense constructs of TrCHId in pDH51 transformation vector.

Figure 130 shows plasmid maps of sense and antisense constructs of TrCHId in pPZP221:35S^2 binary transformation vector.

Figure 131 shows a plasmid map of the cDNA encoding white clover CHRc.

Figure 132 shows the full nucleotide sequence of white clover CHRc cDNA (Sequence ID No: 311).

Figure 133 shows the deduced amino acid sequence of white clover CHRc cDNA (Sequence ID No: 312).

Figure 134 shows plasmid maps of sense and antisense constructs of TrCHRc in pDH51 transformation vector.

Figure 135 shows plasmid maps of sense and antisense constructs of TrCHRc in pPZP221:35S^2 binary transformation vector.

Figure 136 shows a plasmid map of the cDNA encoding white clover CHSa1.

Figure 137 shows the full nucleotide sequence of white clover CHSa1 cDNA (Sequence ID No: 313).

Figure 138 shows the deduced amino acid sequence of white clover CHSa1 cDNA (Sequence ID No: 314).

Figure 139 shows plasmid maps of sense and antisense constructs of TrCHSa1 in pDH51 transformation vector.

Figure 140 shows plasmid maps of sense and antisense constructs of TrCHSa1 in pPZP221:35S^2 binary transformation vector.

Figure 141 shows a plasmid map of the cDNA encoding white clover CHSa3.

Figure 142 shows the full nucleotide sequence of white clover CHSa3 cDNA (Sequence ID No: 315).

Figure 143 shows the deduced amino acid sequence of white clover CHSa3 cDNA (Sequence ID No: 316).

Figure 144 shows plasmid maps of sense and antisense constructs of TrCHSa3 in pDH51 transformation vector.

Figure 145 shows plasmid maps of sense and antisense constructs of TrCHSa3 in pPZP221:35S^2 binary transformation vector.

Figure 146 shows a plasmid map of the cDNA encoding white clover CHSc.

Figure 147 shows the full nucleotide sequence of white clover CHSc cDNA (Sequence ID No: 317).

Figure 148 shows the deduced amino acid sequence of white clover CHSc cDNA (Sequence ID No: 318).

Figure 149 shows plasmid maps of sense and antisense constructs of TrCHSc in pDH51 transformation vector.

Figure 150 shows plasmid maps of sense and antisense constructs of TrCHSc in pPZP221:35S^2 binary transformation vector.

Figure 151 shows a plasmid map of the cDNA encoding white clover CHSd2.

Figure 152 shows the full nucleotide sequence of white clover CHSd2 cDNA (Sequence ID No: 319).

Figure 153 shows the deduced amino acid sequence of white clover CHSd2 cDNA (Sequence ID No: 320).

Figure 154 shows plasmid maps of sense and antisense constructs of TrCHSd2 in pDH51 transformation vector.

Figure 155 shows plasmid maps of sense and antisense constructs of TrCHSd2 in pPZP221:35S^2 binary transformation vector.

Figure 156 shows a plasmid map of the cDNA encoding white clover CHSf.

Figure 157 shows the full nucleotide sequence of white clover CHSf cDNA (Sequence ID No: 321).

Figure 158 shows the deduced amino acid sequence of white clover CHSf cDNA (Sequence ID No: 322).

Figure 159 shows plasmid maps of sense and antisense constructs of TrCHSf in pDH51 transformation vector.

Figure 160 shows plasmid maps of sense and antisense constructs of TrCHSf in pPZP221:35S^2 binary transformation vector.

Figure 161 shows a plasmid map of the cDNA encoding white clover CHSh.

Figure 162 shows the full nucleotide sequence of white clover CHSh cDNA (Sequence ID No: 323).

Figure 163 shows the deduced amino acid sequence of white clover CHSh cDNA (Sequence ID No: 324).

Figure 164 shows plasmid maps of sense and antisense constructs of TrCHSh in pDH51 transformation vector.

Figure 165 shows plasmid maps of sense and antisense constructs of TrCHSh in pPZP221:35S^2 binary transformation vector.

Figure 166 shows a plasmid map of the cDNA encoding white clover DFRd.

Figure 167 shows the full nucleotide sequence of white clover DFRd cDNA (Sequence ID No: 325).

Figure 168 shows the deduced amino acid sequence of white clover DFRd cDNA (Sequence ID No: 326).

Figure 169 shows plasmid maps of sense and antisense constructs of TrDFRd in pDH51 transformation vector.

Figure 170 shows plasmid maps of sense and antisense constructs of TrDFRd in pPZP221:35S^2 binary transformation vector.

Figure 171 shows a plasmid map of the cDNA encoding white clover F3Ha.

Figure 172 shows the full nucleotide sequence of white clover F3Ha cDNA (Sequence ID No: 327).

Figure 173 shows the deduced amino acid sequence of white clover F3Ha cDNA (Sequence ID No: 328).

Figure 174 shows plasmid maps of sense and antisense constructs of TrF3Ha in pDH51 transformation vector.

Figure 175 shows plasmid maps of sense and antisense constructs of TrF3Ha in pPZP221:35S^2 binary transformation vector.

Figure 176 shows a plasmid map of the cDNA encoding white clover PALa.

Figure 177 shows the full nucleotide sequence of white clover PALa cDNA (Sequence ID No: 329).

Figure 178 shows the deduced amino acid sequence of white clover PALa cDNA (Sequence ID No: 330).

Figure 179 shows plasmid maps of sense and antisense constructs of TrPALa in pDH51 transformation vector.

Figure 180 shows plasmid maps of sense and antisense constructs of TrPALa in pPZP221:35S^2 binary transformation vector.

Figure 181 shows a plasmid map of the cDNA encoding white clover PALb.

Figure 182 shows the full nucleotide sequence of white clover PALb cDNA (Sequence ID No: 331).

Figure 183 shows the deduced amino acid sequence of white clover PALb cDNA (Sequence ID No: 332).

Figure 184 shows plasmid maps of sense and antisense constructs of TrPALb in pDH51 transformation vector.

Figure 185 shows plasmid maps of sense and antisense constructs of TrPALb in pPZP221:35S^2 binary transformation vector.

Figure 186 shows a plasmid map of the cDNA encoding white clover PALf.

Figure 187 shows the full nucleotide sequence of white clover PALf cDNA (Sequence ID No: 333).

Figure 188 shows the deduced amino acid sequence of white clover PALf cDNA (Sequence ID No: 334).

Figure 189 shows plasmid maps of sense and antisense constructs of TrPALf in pDH51 transformation vector.

Figure 190 shows plasmid maps of sense and antisense constructs of TrPALf in pPZP221:35S^2 binary transformation vector.

Figure 191 shows a plasmid map of the cDNA encoding white clover VRa.

Figure 192 shows the full nucleotide sequence of white clover VRa cDNA (Sequence ID No: 335).

Figure 193 shows the deduced amino acid sequence of white clover VRa cDNA (Sequence ID No: 336).

Figure 194 shows plasmid maps of sense and antisense constructs of TrVRa in pDH51 transformation vector.

Figure 195 shows plasmid maps of sense and antisense constructs of TrVRa in pPZP221:35S^2 binary transformation vector.

Figure 196 shows A, infiltration of Arabidopsis plants; B, selection of transgenic Arabidopsis plants on medium containing 75 μg/ml gentamycin; C, young transgenic Arabidopsis plants; D, E, two representative results of real-time PCR analysis of Arabidopsis transformed with chimeric genes involved in flavonoid biosynthesis.

Figure 197 shows the genetic map detailing the relation of perennial ryegrass genes involved in flavonoid biosynthesis.

EXAMPLE 1

Preparation of cDNA libraries, isolation and sequencing of cDNAs coding for CHI, CHI-like, CHS, CHS-like, CHR, CHR-like, DFR, DFR-like, LCR, LCR-like, F3'5'H, F3'5'H-like, F3H, F3H-like, F3'H, F3'H-like, PAL, PAL-like, VR and VR-like proteins from white clover (*Trifolium repens*) and perennial ryegrass (*Lolium perenne*)

cDNA libraries representing mRNAs from various organs and tissues of white clover (*Trifolium repens*) and perennial ryegrass (*Lolium perenne*) were prepared. The characteristics of the white clover and perennial ryegrass libraries, respectively, are described below (Tables 1 and 2).

TABLE 1

cDNA libraries from white clover (*Trifolium repens*)

Library	Organ/Tissue
01wc	Whole seedling, light grown
02wc	Nodulated root 3, 5, 10, 14, 21 &28 day old seedling
03wc	Nodules pinched off roots of 42 day old rhizobium inoculated plants
04wc	Cut leaf and stem collected after 0, 1, 4, 6 &14 h after cutting
05wc	Inflorescences: <50% open, not fully open and fully open
06wc	Dark grown etiolated
07wc	Inflorescence – very early stages, stem elongation, < 15 petals, 15-20 petals
08wc	seed frozen at –80°C, imbibed in dark overnight at 10°C
09wc	Drought stressed plants
10wc	AMV infected leaf
11wc	WCMV infected leaf
12wc	Phosphorus starved plants
13wc	Vegetative stolon tip
14wc	stolon root initials
15wc	Senescing stolon
16wc	Senescing leaf

TABLE 2

cDNA libraries from perennial ryegrass (*Lolium perenne*)

Library	Organ/Tissue
01rg	Roots from 3-4 day old light-grown seedlings
02rg	Leaves from 3-4 day old light-grown seedlings
03rg	Etiolated 3-4 day old dark-grown seedlings
04rg	Whole etiolated seedlings (1-5 day old and 17 days old)
05rg	Senescing leaves from mature plants

Library	Organ/Tissue
06rg	Whole etiolated seedlings (1-5 day old and 17 days old)
07rg	Roots from mature plants grown in hydroponic culture
08rg	Senescent leaf tissue
09rg	Whole tillers and sliced leaves (0, 1, 3, 6, 12 and 24 h after harvesting)
10rg	Embryogenic suspension-cultured cells
11rg	Non-embryogenic suspension-cultured cells
12rg	Whole tillers and sliced leaves (0, 1, 3, 6, 12 and 24 h after harvesting)
13rg	Shoot apices including vegetative apical meristems
14rg	Immature inflorescences including different stages of inflorescence meristem and inflorescence development
15rg	Defatted pollen
16rg	Leaf blades and leaf sheaths (*rbcL*, *rbcS*, *cab*, *wir2A* subtracted)
17rg	Senescing leaves and tillers
18rg	Drought-stressed tillers (pseudostems from plants subjected to PEG-simulated drought stress)
19rg	Non-embryogenic suspension-cultured cells subjected to osmotic stress (grown in media with half-strength salts) (1, 2, 3, 4, 5, 6, 24 and 48 h after transfer)
20rg	Non-embryogenic suspension-cultured cells subjected to osmotic stress (grown in media with double-strength salts) (1, 2, 3, 4, 5, 6, 24 and 48 h after transfer)
21rg	Drought-stressed tillers (pseudostems from plants subjected to PEG-simulated drought stress)
22rg	Spikelets with open and maturing florets
23rg	Mature roots (specific subtraction with leaf tissue)

The cDNA libraries may be prepared by any of many methods available. For example, total RNA may be isolated using the Trizol method (Gibco-BRL, USA) or the RNeasy Plant Mini kit (Qiagen, Germany), following the manufacturers' instructions. cDNAs may be generated using the SMART PCR

cDNA synthesis kit (Clontech, USA), cDNAs may be amplified by long distance polymerase chain reaction using the Advantage 2 PCR Enzyme system (Clontech, USA), cDNAs may be cleaned using the GeneClean spin column (Bio 101, USA), tailed and size fractionated, according to the protocol provided by Clontech. The cDNAs may be introduced into the pGEM-T Easy Vector system 1 (Promega, USA) according to the protocol provided by Promega. The cDNAs in the pGEM-T Easy plasmid vector are transfected into *Escherichia coli* Epicurian coli XL10-Gold ultra competent cells (Stratagene, USA) according to the protocol provided by Stratagene.

Alternatively, the cDNAs may be introduced into plasmid vectors for first preparing the cDNA libraries in Uni-ZAP XR vectors according to the manufacturer's protocol (Stratagene Cloning Systems, La Jolla, CA, USA). The Uni-ZAP XR libraries are converted into plasmid libraries according to the protocol provided by Stratagene. Upon conversion, cDNA inserts will be contained in the plasmid vector pBluescript. In addition, the cDNAs may be introduced directly into precut pBluescript II SK(+) vectors (Stratagene) using T4 DNA ligase (New England Biolabs), followed by transfection into *E. coli* DH10B cells according to the manufacturer's protocol (GIBCO BRL Products).

Once the cDNA inserts are in plasmid vectors, plasmid DNAs are prepared from randomly picked bacterial colonies containing recombinant plasmids, or the insert cDNA sequences are amplified via polymerase chain reaction using primers specific for vector sequences flanking the inserted cDNA sequences. Plasmid DNA preparation may be performed robotically using the Qiagen QiaPrep Turbo kit (Qiagen, Germany) according to the protocol provided by Qiagen. Amplified insert DNAs are sequenced in dye-terminator sequencing reactions to generate partial cDNA sequences (expressed sequence tags or "ESTs"). The resulting ESTs are analyzed using an Applied Biosystems ABI 3700 sequence analyser.

EXAMPLE 2

DNA sequence analyses

The cDNA clones encoding CHI, CHI-like, CHS, CHS-like, CHR, CHR-like, DFR, DFR-like, LCR, LCR-like, F3'5'H, F3'5'H-like, F3H, F3H-like, F3'H, F3'H-like, PAL, PAL-like, VR and VR-like proteins were identified by conducting BLAST (Basic Local Alignment Search Tool; Altschul *et al.* (1993) *J. Mol. Biol.* 215:403-410) searches. The cDNA sequences obtained were analysed for similarity to all publicly available DNA sequences contained in the *e*Bioinformatics nucleotide database using the BLASTN algorithm provided by the National Center for Biotechnology Information (NCBI). The DNA sequences were translated in all reading frames and compared for similarity to all publicly available protein sequences contained in the SWISS-PROT protein sequence database using BLASTx algorithm (v 2.0.1) (Gish and States (1993) *Nature Genetics* 3:266-272) provided by the NCBI.

The cDNA sequences obtained and identified were then used to identify additional identical and/or overlapping cDNA sequences generated using the BLASTN algorithm. The identical and/or overlapping sequences were subjected to a multiple alignment using the CLUSTALw algorithm, and to generate a consensus contig sequence derived from this multiple sequence alignment. The consensus contig sequence was then used as a query for a search against the SWISS-PROT protein sequence database using the BLASTx algorithm to confirm the initial identification.

EXAMPLE 3

Identification and full-length sequencing of cDNAs encoding perennial ryegrass F3OH and white clover BANa, CHIa, CHId, CHRc, CHSa1, CHSa3, CHSc, CHSd2, CHSf, CHSh, DFRd, F3Ha, PALa, PALb, PALf and VRa proteins

To fully characterise for the purposes of the generation of probes for hybridisation experiments and the generation of transformation vectors, a set of

cDNAs encoding perennial ryegrass F3OH and white clover BANa, CHIa, CHId, CHRc, CHSa1, CHSa3, CHSc, CHSd2, CHSf, CHSh, DFRd, F3Ha, PALa, PALb, PALf and VRa proteins was identified and fully sequenced.

Full-length cDNAs were identified from our EST sequence database using relevant published sequences (NCBI databank) as queries for BLAST searches. Full-length cDNAs were identified by alignment of the query and hit sequences using Sequencher (Gene Codes Corp., Ann Arbor, MI 48108, USA). The original plasmid was then used to transform chemically competent XL-1 cells (prepared in-house, $CaCl_2$ protocol). After colony PCR (using HotStarTaq, Qiagen) a minimum of three PCR-positive colonies per transformation were picked for initial sequencing with M13F and M13R primers. The resulting sequences were aligned with the original EST sequence using Sequencher to confirm identity and one of the three clones was picked for full-length sequencing, usually the one with the best initial sequencing result.

Sequencing was completed by primer walking, i.e. oligonucleotide primers were designed to the initial sequence and used for further sequencing. In most cases the sequencing could be done from both 5' and 3' end. The sequences of the oligonucleotide primers are shown in Table 2. In some instances, however, an extended poly-A tail necessitated the sequencing of the cDNA to be completed from the 5' end.

Contigs were then assembled in Sequencher. The contigs include the sequences of the SMART primers used to generate the initial cDNA library as well as pGEM-T Easy vector sequence up to the EcoRI cut site both at the 5' and 3' end.

Plasmid maps and the full cDNA sequences of perennial ryegrass F3OH and white clover BANa, CHIa, CHId, CHRc, CHSa1, CHSa3, CHSc, CHSd2, CHSf, CHSh, DFRd, F3Ha, PALa, PALb, PALf and VRa proteins were obtained (Figures 110, 116, 121, 126, 131, 136, 141, 146, 151, 156, 161, 166, 171, 176, 181, 186 and 191).

TABLE 2

List of primers used for sequencing of the full-length cDNAs

gene name	clone ID	sequencing primer	primer sequence (5'>3')
LpF3OH	08rg1YsF07	08rg1YsF07.f1	TTGAGAGCTTCGTCGACC
		08rg1YsF07.r1	AACTCCTCGTAGTACTCC
TrCHRc	11wc1IsD03	11wc1IsD03.f1	TTCAATTGGAGTACTTGG
		11wc1IsD03.r1	ACTCCTTGTTCATATAACC
TrCHSa1	02wc2FsD07	02wc2FsD07.f1	ACATGGTGGTGGTTGAGG
		02wc2FsD07.f2	TGCTGCACTCATTGTTGG
		02wc2FsD07.f3	ACATTGATAAGGCATTGG
TrCHSa3	05wc1RsB06	05wc1RsB06.f1	AGGAGGCTGCAGTCAAGG
		05wc1RsB06.f2	TGCCTGAAATTGAGAAACC
		05wc1RsB06.f3	AAAGCTAGCCTTGAAGCC
TrCHSc	07wc1TsE12	07wc1TsE12.f1	TCGGACATAACTCATGTGG
		07wc1TsE12.f2	TTGGGTTGGAGAATAAGG
		07wc1TsE12.r1	TGGACATTTATTGGTTGC
		07wc1TsE12.r2	TATCATGTCTGGAAATGC
TrCHSd2	07wc1XsD03	07wc1XsD03.f1	TTTATGTGAGTACATGGC
		07wc1XsD03.f2	AGCAGCTGTGATTGTAGG
		07wc1XsD03.f3	TGAGAAAGCTCTTGTTGAGG
TrCHSf	07wc1UsD07	07wc1UsD07.f1	AGATTGCATCAAAGAATGG
		07wc1UsD07.r1	GGTCCAAAAGCCAATCC
TrCHSh	13wc2IsG04	13wc2IsG04.f1	TAAGACGAGACATAGTGG
		13wc2IsG04.r1	TATTCACTAAGCACATGC
TrDFRd	12wc1CsE09	12wc1CsE09.f1	TTACCTCGTCTGTCTCG
		12wc1CsE09.r1	AACACACACATGTCTACC
TrF3Ha	07wc1LsG03	07wc1LsG03.f1	TGAAGGATTGGAGAGAGC
		07wc1LsG03.r1	TACACAGTTGCATCTGG
TrPALa	04wc1UsB03	04wc1UsB03.f1	ATCGGAATCTGCTAGAGC
		04wc1UsB03.f2	TGTTGGTTCTGGTTTAGC
		04wc1UsB03.r1	TTCATATGCAATCCTTGC
		04wc1UsB03.r2	TCTTGGTTGTGTTGTTCC

TrPALb	05wc1PsH02	05wc1PsH02.f1	TGGGACTGATAGTTATGG
		05wc1PsH02.f2	TCTTGCTCTTGTTAATGG
		05wc1PsH02.r1	AGCACCATTCCACTCTCC
		05wc1PsH02.r2	TTCTCTTCGCTACTTGGC
TrPALf	13wc2AsD12	13wc2AsD12.f1	ATAGTGGTGTGAGGGTGG
		13wc2AsD12.f2	TCTTGTTAATGGTACTGC
		13wc2AsD12.r1	ATTTATCGCACTCTTCGC
		13wc2AsD12.r2	AAAGTGGAAGACATGAGC
TrVRa	11wc1NsA07	11wc1NsA07.f1	AAGAACAGTGGATGGAGC
		11wc1NsA07.r1	TCAACTCATCTACTGATAG

EXAMPLE 4

Development of transformation vectors containing chimeric genes with cDNA sequences from perennial ryegrass F3OH and white clover BANa, CHIa, CHId, CHRc, CHSa1, CHSa3, CHSc, CHSd2, CHSf, CHSh, DFRd, F3Ha, PALa, PALb, PALf and VRa

To alter the expression of the proteins involved in flavonoid biosynthesis, protein binding, metal chelation, anti-oxidation, UV-light absorption, tolerance to biotic stresses such as viruses, micro-organisms, insects and fungal pathogens; pigmentation in for example flowers and leaves; herbage quality and bloat-safety and isoflavonoid content leading to health benefits, perennial ryegrass F3OH and white clover BANa, CHIa, CHId, CHRc, CHSa1, CHSa3, CHSc, CHSd2, CHSf, CHSh, DFRd, F3Ha, PALa, PALb, PALf and VRa, through antisense and/or sense suppression technology and for over-expression of these key enzymes in transgenic plants, a set of sense and antisense transformation vectors was produced.

cDNA fragments were generated by high fidelity PCR using the original pGEM-T Easy plasmid cDNA as a template. The primers used (Table 3) contained recognition sites for appropriate restriction enzymes, for example EcoRI and XbaI, for directional and non-directional cloning into the target vector. After PCR amplification and restriction digest with the appropriate restriction enzyme (usually

XbaI), the cDNA fragments were cloned into the corresponding site in pDH51, a pUC18-based transformation vector containing a CaMV 35S expression cassette. The orientation of the constructs (sense or antisense) was checked by DNA sequencing through the multi-cloning site of the vector. Transformation vectors containing chimeric genes using full-length open reading frame cDNAs encoding perennial ryegrass F3OH and white clover BANa, CHIa, CHId, CHRc, CHSa1, CHSa3, CHSc, CHSd2, CHSf, CHSh, DFRd, F3Ha, PALa, PALb, PALf and VRa proteins in sense and antisense orientations under the control of the CaMV 35S promoter were generated (Figures 113, 119, 124, 129, 134, 139, 144, 149, 154, 159, 164, 169, 174, 179, 184, 189 and 194).

TABLE 3

List of primers used to PCR-amplify the open reading frames

gene name	clone ID	primer	primer sequence (5'->3')
LpF3OH	08rg1YsF07	08rg1YsF07f	GAATTCTAGAAGCAGAAAGTACGGACATCAGC
		08rg1YsF07r	GAATTCTAGAACCATATGGCGACACATCG
TrBANa	05wc2XsG02	05wc2XsG02f	GGATCCTCTAGAGCACTAGTGTGTATAAGTTTCTTGG
		05wc2XsG02r	GGATCCTCTAGACCCCCTTAGTCTTAAAATACTCG
TrCHIa	06wc2AsF12	06wc2AsF12f	GAATTCTAGAGATCTGAAACAACATAGTCACC
		06wc2AsF12r	GAATTCTAGATCAATCTTGTGCTGCAATGC
TrCHId	12wc1FsG04	12wc1FsG04f	GAATTCTAGAAAGTTCAACGAGATCAATGG
		12wc1FsG04r	GAATTCTAGATTCCGCTTGGTCTTTATTGC
TrCHRc	11wc1IsD03	11wc1IsD03f	GAATTCTAGAACATGGGTAGTGTTGAAATTCC
		11wc1IsD03r	GAATTCTAGAAGATATTGAGTGAGCTTAAGG
TrCHSa1	02wc2FsD07	02wc2FsD07f	GACGTCGACATTACATACATAGCAGGAAC
		02wc2FsD07r	GACGTCGACAGTCTCTCATTCTCATATAGC
TrCHSa3	05wc1RsB06	05wc1RsB06f	GAATTCTAGAAGATATGGTGAGTGTAGCTG
		05wc1RsB06r	GAATTCTAGAATCACACATCTTATATAGCC
TrCHSc	07wc1TsE12	07wc1TsE12f	GAATTCTAGAAGAAGAAATATGGGAGACGAAGG
		07wc1TsE12r	GAATTCTAGAAAGACTTCATGCACACAAGTTCC
TrCHSd2	07wc1XsD03	07wc1XsD03f	GAATTCTAGAATAACCTATCAGTACTCACC
		07wc1XsD03r	GAATTCTAGAATCTAGGCAATTTAAGTGGC

TrCHSf	07wc1UsD07	07wc1UsD07f	GAATTCTAGATGATTCATTGTTTGTTTCCATAAC
		07wc1UsD07r	GAATTCTAGAACATATTCATCTTCCTATCAC
TrCHSh	13wc2IsG04	13wc2IsG04f	GAATTCTAGATCCAAATTCTCGTACCTCACC
		13wc2IsG04r	GAATTCTAGATAGTTCACATCTCTCGGCAGG
TrDFRd	12wc1CsE09	12wc1CsE09f	GACGTCGACACAACAGTCTTCCACTTGAGC
		12wc1CsE09r	GACGTCGACTCTATACTCTGGTAACTATAGG
TrF3Ha	07wc1LsG03	07wc1LsG03f	GAATTCTAGAACCACACAACACACAAACACC
		07wc1LsG03r	GAATTCTAGAACCAAGCAGCTTAATACACG
TrPALa	04wc1UsB03	04wc1UsB03f	AGTACTGCAGAGATATGGAAGTAGTAGCAGCAGC
		04wc1UsB03r	AGTACTGCAGTAGCAAACCAGTTCCCAACTCC
TrPALb	05wc1PsH02	05wc1PsH02f	AGTACTGCAGATAATGGAGGGAATTACCAATGG
		05wc1PsH02r	AGTACTGCAGTGCTAATTAACATATTGGTAGAGG
TrPALf	13wc2AsD12	13wc2AsD12f	AGTACTGCAGATAATGGAGGGAATTACCAATGG
		13wc2AsD12r	AGTACTGCAGTGCTAATTAACATATTGGTAGAGG
TrVRa	11wc1NsA07	11wc1NsA07f	AGTACTGCAGATAAAGAGAGTCAAAAATGGC
		11wc1NsA07r	AGTACTGCAGAACACATACTTAGAGATAGCC

EXAMPLE 5

Development of binary transformation vectors containing chimeric genes with cDNA sequences from perennial ryegrass F3OH and white clover BANa, CHIa, CHId, CHRc, CHSa1, CHSa3, CHSc, CHSd2, CHSf, CHSh, DFRd, F3Ha, PALa, PALb, PALf and VRa

To alter the expression of the proteins involved in flavonoid biosynthesis, protein binding, metal chelation, anti-oxidation, UV-light absorption, tolerance to biotic stresses such as viruses, micro-organisms, insects and fungal pathogens; pigmentation in for example flowers and leaves; herbage quality and bloat-safety and isoflavonoid content leading to health benefits, perennial ryegrass F3OH and white clover BANa, CHIa, CHId, CHRc, CHSa1, CHSa3, CHSc, CHSd2, CHSf, CHSh, DFRd, F3Ha, PALa, PALb, PALf and VRa, through antisense and/or sense suppression technology and for over-expression of these key proteins in transgenic plants, a set of sense and antisense binary transformation vectors was produced.

cDNA fragments were generated by high fidelity PCR using the original pGEM-T Easy plasmid cDNA as a template. The primers used (Table 3) contained recognition sites for appropriate restriction enzymes, for example EcoRI and XbaI, for directional and non-directional cloning into the target vector. After PCR amplification and restriction digest with the appropriate restriction enzyme (usually XbaI), the cDNA fragments were cloned into the corresponding site in a modified pPZP binary vector (Hajdukiewicz *et al.*, 1994). The pPZP221 vector was modified to contain the $35S^2$ cassette from pKYLX71:$35S^2$ as follows. pKYLX71:$35S^2$ was cut with ClaI. The 5' overhang was filled in using Klenow and the blunt end was A-tailed with Taq polymerase. After cutting with EcoRI, the 2kb fragment with an EcoRI-compatible and a 3'-A tail was gel-purified. pPZP221 was cut with HindIII and the resulting 5' overhang filled in and T-tailed with Taq polymerase. The remainder of the original pPZP221 multi-cloning site was removed by digestion with EcoRI, and the expression cassette cloned into the EcoRI site and the 3' T overhang restoring the HindIII site. This binary vector contains between the left and right border the plant selectable marker gene aaaC1 under the control of the 35S promoter and 35S terminator and the pKYLX71:$35S^2$-derived expression cassette with a CaMV 35S promoter with a duplicated enhancer region and an rbcS terminator.

The orientation of the constructs (sense or antisense) was checked by restriction enzyme digest. Transformation vectors containing chimeric genes using full-length open reading frame cDNAs encoding perennial ryegrass F3OH and white clover BANa, CHIa, CHId, CHRc, CHSa1, CHSa3, CHSc, CHSd2, CHSf, CHSh, DFRd, F3Ha, PALa, PALb, PALf and VRa proteins in sense and antisense orientations under the control of the CaMV $35S^2$ promoter were generated (Figures 114, 120, 125, 130, 135, 140, 145, 150, 155, 160, 165, 170, 175, 180, 185, 190 and 195).

EXAMPLE 6

Production and analysis of transgenic Arabidopsis plants carrying chimeric perennial ryegrass F3OH and white clover BANa, CHIa, CHId, CHRc, CHSa1, CHSa3, CHSc, CHSd2, CHSf, CHSh, DFRd, F3Ha, PALa, PALb, PALf and VRa genes involved in flavonoid biosynthesis

A set of transgenic Arabidopsis plants carrying chimeric perennial ryegrass and white clover genes involved in flavonoid biosynthesis, protein binding, metal chelation, anti-oxidation, UV-light absorption, tolerance to biotic stresses such as viruses, micro-organisms, insects and fungal pathogens; pigmentation in for example flowers and leaves; herbage quality and bloat-safety and isoflavonoid content leading to health benefits, were produced.

pPZP221-based transformation vectors with *Lp*F3OH and *Tr*BANa, *Tr*CHIa, *Tr*CHId, *Tr*CHRc, *Tr*CHSa1, *Tr*CHSa3, *Tr*CHSc, *Tr*CHSd2, *Tr*CHSf, *Tr*CHSh, *Tr*DFRd, *Tr*F3Ha, *Tr*PALa, *Tr*PALb, *Tr*PALf and *Tr*VRa cDNAs comprising the full open reading frame sequences in sense and antisense orientations under the control of the CaMV 35S promoter with duplicated enhancer region ($35S^2$) were generated as detailed in Example 6.

Agrobacterium-mediated gene transfer experiments were performed using these transformation vectors.

The production of transgenic Arabidopsis plants carrying the perennial ryegrass F3OH and white clover BANa, CHIa, CHId, CHRc, CHSa1, CHSa3, CHSc, CHSd2, CHSf, CHSh, DFRd, F3Ha, PALa, PALb, PALf and VRa cDNAs under the control of the CaMV 35S promoter with duplicated enhancer region ($35S^2$) is described here in detail.

Preparation of Arabidopsis plants

Seedling punnets were filled with Debco seed raising mixture (Debco Pty. Ltd.) to form a mound. The mound was covered with two layers of anti-bird netting secured with rubber bands on each side. The soil was saturated with water and

enough seeds (*Arabidopsis thaliana* ecotype Columbia, Lehle Seeds #WT-02) sown to obtain approximately 15 plants per punnet. The seeds were then vernalised by placing the punnets at 4 ºC. After 48 hours the punnets were transferred to a growth room at 22 ºC under fluorescent light (constant illumination, 55 $\mu molm^{-2}s^{-1}$) and fed with Miracle-Gro (Scotts Australia Pty. Ltd.) once a week. Primary bolts were removed as soon as they appeared. After 4 – 6 days the secondary bolts were approximately 6 cm tall, and the plants were ready for vacuum infiltration.

Preparation of *Agrobacterium*

Agrobacterium tumefaciens strain AGL-1 were streaked on LB medium containing 50 μg/ml rifampicin and 50 μg/ml kanamycin and grown at 27 ºC for 48 hours. A single colony was used to inoculate 5 ml of LB medium containing 50 μg/ml rifampicin and 50 μg/ml kanamycin and grown over night at 27 ºC and 250 rpm on an orbital shaker. The overnight culture was used as an inoculum for 500 ml of LB medium containing 50 μg/ml kanamycin only. Incubation was over night at 27 ºC and 250 rpm on an orbital shaker in a 2 l Erlenmeyer flask.

The overnight cultures were centrifuged for 15 min at 5500 xg and the supernatant discarded. The cells were resuspended in 1 *l* of infiltration medium [5% (w/v) sucrose, 0.03% (v/v) Silwet-L77 (Vac-In-Stuff, Lehle Seeds #VIS-01)] and immediately used for infiltration.

Vacuum infiltration

The *Agrobacterium* suspension was poured into a container (Décor Tellfresh storer, #024) and the container placed inside the vacuum desiccator (Bel Art, #42020-0000). A punnet with *Arabidopsis* plants was inverted and dipped into the Agrobacterium suspension and a gentle vacuum (250 mm Hg) was applied for 2 min. After infiltration, the plants were returned to the growth room where they were kept away from direct light overnight. The next day the plants were returned to full direct light and allowed to grow until the siliques were fully developed. The plants were then allowed to dry out, the seed collected from the siliques and either

stored at room temperature in a dry container or used for selection of transformants.

Selection of transformants

Prior to plating the seeds were sterilised as follows. Sufficient seeds for one 150 mm petri dish (approximately 40 mg or 2000 seeds) were placed in a 1.5 ml microfuge tube. 500 μl 70% ethanol were added for 2 min and replaced by 500 μl sterilisation solution (H_2O:4% chlorine:5% SDS, 15:8:1). After vigorous shaking, the tube was left for 10 min after which time the sterilisation solution was replaced with 500 μl sterile water. The tube was shaken and spun for 5 sec to sediment the seeds. The washing step was repeated 3 times and the seeds were left covered with approximately 200 μl sterile water.

The seeds were then evenly spread on 150 mm petri dishes containing germination medium (4.61 g Murashige & Skoog salts, 10 g sucrose, 1 ml 1 M KOH, 2 g Phytagel, 0.5 g MES and 1 ml 1000x Gamborg's B-5 vitamins per litre) supplemented with 250 μg/ml timetin and 75 μg/ml gentamycin. After vernalisation for 48 hours at 4 ºC the plants were grown under continuous fluorescent light (55 μmol m-2s-1) at 22 ºC to the 6 – 8 leaf stage and transferred to soil.

Preparation of genomic DNA

3 – 4 leaves of Arabidopsis plants regenerated on selective medium were harvested and freeze-dried. The tissue was homogenised on a Retsch MM300 mixer mill, then centrifuged for 10 min at 1700xg to collect cell debris. Genomic DNA was isolated from the supernatant using Wizard Magnetic 96 DNA Plant System kits (Promega) on a Biomek FX (Beckman Coulter). 5 μl of the sample (50 μl) were then analysed on an agarose gel to check the yield and the quality of the genomic DNA.

Analysis of DNA using real-time PCR

Genomic DNA was analysed for the presence of the transgene by real-time PCR using SYBR Green chemistry. PCR primer pairs (Table 4) were designed

using MacVector (Accelrys). The forward primer was located within the $35S^2$ promoter region and the reverse primer within the transgene to amplify products of approximately 150 - 250 bp as recommended. The positioning of the forward primer within the $35S^2$ promoter region guaranteed that homologous genes in Arabidopsis were not detected.

5 μl of each genomic DNA sample was run in a 50 μl PCR reaction including SYBR Green on an ABI (Applied Biosystems) together with samples containing DNA isolated from wild type Arabidopsis plants (negative control), samples containing buffer instead of DNA (buffer control) and samples containing the plasmid used for transformation (positive plasmid control).

Plants were obtained after transformation with all chimeric constructs and selection on medium containing gentamycin. The selection process and two representative real-time PCR analyses are shown in Figure xx.

TABLE 4

List of primers used for Real-time PCR analysis of Arabidopsis plants transformed with chimeric perennial ryegrass genes involved in flavonoid biosynthesis

construct	primer 1 (forward)	primer 2 (reverse)
pPZP221LpF3OHsense	TTGGAGAGGACACGCTGAAATC	AGGAGAGGGTTGGACATCGC
pPZP221LpF3OHanti	CATTTCATTTGGAGAGGACACGC	ACGAGGAGTTCTGGAAGATGGG
pPZP221TrBANasense	TTGGAGAGGACACGCTGAAATC	GCAACAAAACCAGTGCCACC
pPZP221TrBANaanti	TCATTTGGAGAGGACACGCTG	GATGATTGCCCCAGCAAGG
pPZP221TrCHIasense	CATTTCATTTGGAGAGGACACGC	CAAGGTTCTCGACTTGGATTGC
pPZP221TrCHIaanti	TCATTTGGAGAGGACACGCTG	AGATTACCTGCCTTGTTGAACGAG
pPZP221TrCHIdsense	TCATTTGGAGAGGACACGCTG	GACGGTAGGAGGGAATAGATTGTTC
pPZP221TrCHIdanti	TCATTTGGAGAGGACACGCTG	CCAGGTTATCCGAGTTATTCAACG
pPZP221TrCHRcsense	CCACTATCCTTCGCAAGACCC	TCCCATTCCAACCACAGGC
pPZP221TrCHRcanti	TCATTTGGAGAGGACACGCTG	CAAGCCAGGACTCAGTGACCTATG
pPZP221TrCHSa1sense	TCATTTGGAGAGGACACGCTG	CTGGTCAACACGATTTGCTGG
pPZP221TrCHSa1anti	TCATTTGGAGAGGACACGCTG	AACCACAGGAGAAGGACTTGACTG

pPZP221TrCHSa3sense	CATTTCATTTGGAGAGGACACGC	AACACGGTTTGGTGGATTTGC
pPZP221TrCHSa3anti	TCATTTGGAGAGGACACGCTG	ACAACTGGAGAAGGACTTGATTGG
pPZP221TrCHScsense	TTGGAGAGGACACGCTGAAATC	ACAAGTTGGTGAGGGAATGCC
pPZP221TrCHScanti	TCATTTGGAGAGGACACGCTG	GGGATTGATACTTGCTTTTGGACC
pPZP221TrCHSd2sense	CCCACTATCCTTCGCAAGACC	AGTTGCAGTGCCGATTGCC
pPZP221TrCHSd2anti	CATTTCATTTGGAGAGGACACGC	AAGATGGACTTGCCACAACAGG
pPZP221TrCHSfsense	CATTTCATTTGGAGAGGACACGC	TCGTTGCCTTTCCCTGAGTAGG
pPZP221TrCHSfanti	TCATTTGGAGAGGACACGCTG	GATTGGCTTTTGGACCAGGG
pPZP221TrCHShsense	TCATTTGGAGAGGACACGCTG	CGGTCACCATTTTTTTGTTGGAGG
pPZP221TrCHShanti	TCATTTGGAGAGGACACGCTG	TGTTGTTTGGGTTTGGACCG
pPZP221TrDFRdsense	CATTTCATTTGGAGAGGACACGC	ATTGAGATTTTGGACGGTGGC
pPZP221TrDFRdanti	CATTTCATTTGGAGAGGACACGC	CGCAACCTGGATTGTTGAGAGC
pPZP221TrF3Hasense	TCATTTGGAGAGGACACGCTG	TCTTCCCTAACGAAACTTGACTCG
pPZP221TrF3Haanti	TCATTTGGAGAGGACACGCTG	GAACAACAACTTAGGGACTTGGAGG
pPZP221TrPALasense	ATGACGCACAATCCCACTATCC	TTGCCTCAGCAGCCACACC
pPZP221TrPALaanti	GGAGAGGACACGCTGAAATCAC	TGCCAAAAGAGGTTGAAAGTGC
pPZP221TrPALbsense	ATCCCACTATCCTTCGCAAGACCC	AATGACTCCCCATCAACGACTCCG
pPZP221TrPALbanti	TTGGAGAGGACACGCTGAAATC	GACAAATTGTTCACAGCTATGTGCC
pPZP221TrPALfsense	ATCCCACTATCCTTCGCAAGACCC	CACCATACGCTTCACCTCATCC
pPZP221TrPALfanti	TCATTTGGAGAGGACACGCTG	TTGTTAGAGAGGAGTTAGGAACCGC
pPZP221TrVRasense	CCACTATCCTTCGCAAGACCC	GCTTACATCCCTCTTACGTTCTGG
pPZP221TrVRaanti	CCACTATCCTTCGCAAGACCC	AAAAGCTCGTGGACGCTGG

EXAMPLE 7

Genetic mapping of perennial ryegrass genes involved in flavonoid biosynthesis, protein binding, metal chelation, anti-oxidation, UV-light absorption, tolerance to biotic stresses such as viruses, micro-organisms, insects and fungal pathogens; pigmentation in for example flowers and leaves; herbage quality and bloat-safety and isoflavonoid content leading to health benefits

The cDNAs representing genes involved in flavonoid biosynthesis, protein binding, metal chelation, anti-oxidation, UV-light absorption, tolerance to biotic stresses such as viruses, micro-organisms, insects and fungal pathogens;

pigmentation in for example flowers and leaves; herbage quality and bloat-safety and isoflavonoid content leading to health benefits, were amplified by PCR from their respective plasmids, gel-purified and radio-labelled for use as probes to detect restriction fragment length polymorphisms (RFLPs). RFLPs were mapped in the F_1 (first generation) population, NA_6 x AU_6. This population was made by crossing an individual (NA_6) from a North African ecotype with an individual (AU_6) from the cultivar Aurora, which is derived from a Swiss ecotype. Genomic DNA of the 2 parents and 114 progeny was extracted using the 1 x CTAB method of Fulton et al. (1995).

Probes were screened for their ability to detect polymorphism using the DNA (10 μg) of both parents and 5 F_1 progeny restricted with the enzymes DraI, EcoRI, EcoRV or HindIII. Hybridisations were carried out using the method of Sharp et al. (1988). Polymorphic probes were screened on a progeny set of 114 individuals restricted with the appropriate enzyme (Figure 115).

RFLP bands segregating within the population were scored and the data was entered into an Excel spreadsheet. Alleles showing the expected 1:1 ratio were mapped using MAPMAKER 3.0 (Lander et al. 1987). Alleles segregating from, and unique to, each parent, were mapped separately to give two different linkage maps. Markers were grouped into linkage groups at a LOD of 5.0 and ordered within each linkage group using a LOD threshold of 2.0.

Loci representing genes involved in flavonoid biosynthesis mapped to the linkage groups as indicated in Table 5 and in Figure 197. These gene locations can now be used as candidate genes for quantitative trait loci associated with flavonoid biosynthesis, protein binding, metal chelation, anti-oxidation, UV-light absorption, tolerance to biotic stresses such as viruses, micro-organisms, insects and fungal pathogens; pigmentation in for example flowers and leaves; herbage quality and bloat-safety and isoflavonoid content leading to health benefits.

TABLE 5

Map locations of ryegrass genes involved in flavonoid biosynthesis across two genetic linkage maps of perennial ryegrass

Probe	Polymorphic	Mapped with	Locus	Linkage group	
				NA_6	AU_6
*Lp*DFRb	Y	*Hind* III	*Lp*DFRb	6	6

REFERENCES

An, G., Watson, B.D., Stachel, S., Gordon, M.P., Nester, E.W. (1985) New cloning vehicles for transformation of higher plants. The EMBO Journal **4**, 227-284

Feinberg, A.P., Vogelstein, B. (1984). A technique for radiolabelling DNA restriction endonuclease fragments to high specific activity. Anal. Biochem. **132**: 6-13.

Frohman *et al.* (1988) Rapid production of full-length cDNAs from rare transcripts: amplification using a single gene-specific oligonucleotide primer. *Proc. Natl. Acad Sci. USA* **85**:8998

Gish and States (1993) Identification of protein coding regions by database similarity search. *Nature Genetics* **3**:266-272

Lander, E.S., Green P., Abrahamson, J., Barlow, A., Daly, M.J., Lincoln, S.E., Newburg, L. (1987). MAPMAKER: an interactive computer package for constructing primary linkage maps of experimental and natural populations. Genomics **1**: 174-181.

Loh, E.Y., Elliott, J.F., Cwirla, S., Lanier, L.L., Davis, M.M. (1989). Polymerase chain reaction with single-sided specificity: Analysis of T-cell receptor delta

chain. *Science* **243**:217-220

Ohara, O., Dorit, R.L., Gilbert, W. (1989). One-sided polymerase chain reaction: The amplification of cDNA. *Proc. Natl. Acad Sci USA* **86**:5673-5677

Sambrook, J., Fritsch, E.F., Maniatis, T. (1989). Molecular Cloning. A Laboratory Manual. Cold Spring Harbour Laboratory Press

Schardl, C.L., Byrd, A.D., Benzion, G., Altschuler, M.A., Hildebrand, D.F., Hunt, A.G. (1987) Design and construction of a versatile system for the expression of foreign genes in plants. Gene **61**, 1-11

Sharp, P.J., Kreis, M., Shewry, P.R., Gale, M.D. (1988). Location of α-amylase sequences in wheat and its relatives. Theor. Appl. Genet. **75**:286-290.

Finally, it is to be understood that various alterations, modifications and/or additions may be made without departing from the spirit of the present invention as outlined herein.

It will also be understood that the term "comprises" (or its grammatical variants) as used in this specification is equivalent to the term "includes" and should not be taken as excluding the presence of other elements or features.

Documents cited in this specification are for reference purposes only and their inclusion is not acknowledgment that they form part of the common general knowledge in the relevant art.

CLAIMS

1. A substantially purified or isolated nucleic acid or nucleic acid fragment encoding a flavonoid biosynthetic enzyme selected from the group consisting of chalcone isomerase (CHI), chalcone synthase (CHS), chalcone reductase (CHR), dihydroflavonol 4-reductase (DFR), leucoanthocyanidin reductase (LCR), flavonoid 3', 5' hydrolase (F3'5'H), flavanone 3-hydroxylase (F3H), flavonoid 3'-hydroxylase (F3'H), phenylalanine ammonia-olyase (PAL) and vestitone reductase (VR) from a clover (*Trifolium*), medic (*Medicago*), ryegrass (*Lolium*) or fescue (*Festuca*) species, or a functionally active fragment or variant thereof.

2. A nucleic acid or nucleic acid fragment according to Claim 1, wherein said clover species is white clover (*Trifolium repens*) and said ryegrass species is perennial ryegrass (*Lolium perenne*).

3. A nucleic acid or nucleic acid fragment according to Claim 1, encoding a CHI or CHI-like protein and including a nucleotide sequence selected from the group consisting of (a) sequences shown in Figures 1, 3, 4, 6, 7, 9, 10, 12, 122 and 127 hereto (Sequence ID Nos: 1, 3 to 7, 8, 10 to 12, 13, 15 and 16, 17, 19 to 22, 307, and 309, respectively); (b) complements of the sequences recited in (a); (c) sequences antisense to the sequences recited in (a) and (b); and (d) functionally active fragments and variants of the sequences recited in (a), (b) and (c).

4. A nucleic acid or nucleic acid fragment according to Claim 1, encoding a CHS or CHS-like protein and including a nucleotide sequence selected from the group consisting of (a) sequences shown in Figures 13, 15, 16, 18, 19, 21, 22, 24, 25, 27, 28, 30, 31, 33, 34, 137, 142, 147, 152, 157 and 162 hereto (Sequence ID Nos: 23, 25 to 63, 64, 66 to 68, 69, 71 to 77, 78, 80 to 90, 91, 93 and 94, 95, 97 to 100, 101, 103 to 105, 106, 313, 315, 317, 319, 321, and 323, respectively); (b) complements of the sequences recited in (a); (c) sequences antisense to the sequences recited in (a) and (b); and (d) functionally active fragments and variants of the sequences recited in (a), (b) and (c).

5. A nucleic acid or nucleic acid fragment according to Claim 1, encoding a CHR or CHR-like protein and including nucleotide sequence selected from the group consisting of (a) sequences shown in Figures 36, 38, 40, 41, 43 and 132 hereto (Sequence ID Nos: 108, 110, 112 to 116, 117, 119 to 134, and 311, respectively); (b) complements of the sequences recited in (a); (c) sequences antisense to the sequences recited in (a) and (b); and (d) functionally active fragments and variants of the sequences recited in (a), (b) and (c).

6. A nucleic acid or nucleic acid fragment according to Claim 1, encoding a DFR or DFR-like protein and including a nucleotide sequence selected from the group consisting of (a) sequences shown in Figures 44, 46, 47, 49, 50, 52, 54, 55, 57, 59, 61, 62, 64, 101, 103, 104, 106, 117 and 167 hereto (Sequence ID Nos: 135, 137 to 146, 147, 149 to 152, 153, 155, 157 and 158, 159, 161, 163, 165 to 167, 168, 170 to 184, 286, 288 to 292, 293, 295 to 297, 305, and 325, respectively); (b) complements of the sequences recited in (a); (c) sequences antisense to the sequences recited in (a) and (b); and (d) functionally active fragments and variants of the sequences recited in (a), (b) and (c).

7. A nucleic acid or nucleic acid fragment according to Claim 1, encoding an LCR or LCR-like protein and including a nucleotide sequence selected from the group consisting of (a) sequences shown in Figures 65 and 67 hereto (Sequence ID Nos: 185 and 187 to 193, respectively); (b) complements of the sequences recited in (a); (c) sequences antisense to the sequences recited in (a) and (b); and (d) functionally active fragments and variants of the sequences recited in (a), (b) and (c).

8. A nucleic acid or nucleic acid fragment according to Claim 1, encoding a F3'5'H or F3'5'H-like protein and including a nucleotide sequence selected from the group consisting of (a) sequences shown in Figures 68, 70 and 72 hereto (Sequence ID Nos: 194, 196, and 198 to 201, respectively); (b) complements of the sequences recited in (a); (c) sequences antisense to the sequences recited in (a) and (b); and (d) functionally active fragments and variants of the sequences recited in (a), (b) and (c).

9. A nucleic acid or nucleic acid fragment according to Claim 1, encoding a F3H or F3H-like protein and including a nucleotide sequence selected from the group consisting of (a) sequences shown in Figures 73, 75, 76, 78, 107, 109, 111 and 172 hereto (Sequence ID Nos: 202, 204 to 244, 245, 247, 298, 300 to 302, 303, and 327, respectively); (b) complements of the sequences recited in (a); (c) sequences antisense to the sequences recited in (a) and (b); and (d) functionally active fragments and variants of the sequences recited in (a), (b) and (c).

10. A nucleic acid or nucleic acid fragment according to Claim 1, encoding a F3'H or F3'H-like protein and including a nucleotide sequence selected from the group consisting of (a) sequences shown in Figures 80 and 82 hereto (Sequence ID Nos: 249, and 251 and 252, respectively); (b) complements of the sequences recited in (a); (c) sequences antisense to the sequences recited in (a) and (b); and (d) functionally active fragments and variants of the sequences recited in (a), (b) and (c).

11. A nucleic acid or nucleic acid fragment according to Claim 1, encoding a PAL or PAL-like protein and including a nucleotide sequence selected from the group consisting of (a) sequences shown in Figures 83, 85, 86, 88, 89, 91, 93, 95, 97, 177, 182 and 187 hereto (Sequence ID Nos: 253, 255 to 257, 258, 260 to 267, 268, 270, 272, 274, 276 and 277, 329, 331, and 333, respectively); (b) complements of the sequences recited in (a); (c) sequences antisense to the sequences recited in (a) and (b); and (d) functionally active fragments and variants of the sequences recited in (a), (b) and (c).

12. A nucleic acid or nucleic acid fragment according to Claim 1, encoding a VR or VR-like protein and including a nucleotide sequence selected from the group consisting of (a) sequences shown in Figures 98, 100 and 192 hereto (Sequence ID Nos: 278, 280 to 285, and 335, respectively); (b) complements of the sequences recited in (a); (c) sequences antisense to the sequences recited in (a) and (b); and (d) functionally active fragments and variants of the sequences recited in (a), (b) and (c).

13. A construct including a nucleic acid or nucleic acid fragment according to Claim 1.

14. A vector including a nucleic acid or nucleic acid fragment according to Claim 1.

15. A vector according to Claim 14, further including a promoter and a terminator, said promoter, nucleic acid or nucleic acid fragment and terminator being operatively linked.

16. A plant cell, plant, plant seed or other plant part, including a construct according to claim 13 or a vector according to Claim 14.

17. A plant, plant seed or other plant part derived from a plant cell or plant according to Claim 16.

18. A method of modifying flavonoid biosynthesis in a plant, said method including introducing into said plant an effective amount of a nucleic acid or nucleic acid fragment according to Claim 1, a construct according to claim 13 and/or a vector according to Claim 14.

19. A method of modifying protein binding, metal chelation, anti-oxidation, and/or UV-light absorption in a plant, said method including introducing into said plant an effective amount of a nucleic acid or nucleic acid fragment according to Claim 1, a construct according to claim 13 and/or a vector according to Claim 14.

20. A method of modifying pigment production in a plant, said method including introducing into said plant an effective amount of a nucleic acid or nucleic acid fragment according to Claim 1, a construct according to claim 13 and/or a vector according to Claim 14.

21. A method of modifying plant defense to a biotic stress, said method including introducing into said plant an effective amount of a nucleic acid or nucleic

acid fragment according to Claim 1, a construct according to claim 13 and/or a vector according to Claim 14.

22. A method according to claim 21 wherein said biotic stress is selected from the group consisting of viruses, microorganisms, insects and fungal pathogens. .

23. A method of modifying forage quality of a plant by disrupting protein foam and/or conferring protection from rumen pasture bloat, said method including introducing into said plant an effective amount of a nucleic acid or nucleic acid fragment according to Claim 1, a construct according to claim 13 and/or a vector according to Claim 14.

24. Use of a nucleic acid or nucleic acid fragment according to Claim 1, and/or nucleotide sequence information thereof, and/or single nucleotide polymorphisms thereof as a molecular genetic marker.

25. A substantially purified or isolated polypeptide from a clover (*Trifolium*), medic (*Medicago*), ryegrass (*Lolium*) or fescue (*Festuca*) species, selected from the group consisting of ASR and ASR-like, A22 and A22-like, CYS and CYS-like, LEA and LEA-like, DHN and DHN-like and PKABA and PKABA-like; and functionally active fragments and variants thereof.

26. A polypeptide according to Claim 25, wherein said clover species is white clover (*Trifolium repens*) and said ryegrass species is perennial ryegrass (*Lolium perenne*).

27. A polypeptide according to Claim 25, wherein said polypeptide is CHI or CHI-like and includes an amino acid sequence selected from the group consisting of sequences shown in Figures 2, 5, 8, 11, 123 and 128 hereto (Sequence ID Nos: 2, 9, 14, 18, 308, and 310, respectively); and functionally active fragments and variants thereof.

28. A polypeptide according to Claim 25, wherein said polypeptide is CHS or CHS-like and includes an amino acid sequence selected from the group

consisting of sequences shown in Figures 14, 17, 20, 23, 26, 29, 32, 35, 138, 143, 148, 153, 158 and 163 hereto (Sequence ID Nos: 24, 65, 70, 79, 92, 96, 102, 107, 314, 316, 318, 320, 322, and 324, respectively); and functionally active fragments and variants thereof.

29. A polypeptide according to Claim 25, wherein said polypeptide is CHR or CHR-like and includes an amino acid sequence selected from the group consisting of sequences shown in Figures 37, 39, 42 and 133 hereto (Sequence ID Nos: 109, 111, 118, and 312, respectively); and functionally active fragments and variants thereof.

30. A polypeptide according to Claim 25, wherein said polypeptide is DFR or DFR-like and includes an amino acid sequence selected from the group consisting of sequences shown in Figures 45, 48, 51, 53, 56, 58, 60, 63, 102, 105, 118 and 168 hereto (Sequence ID Nos: 136, 148, 54, 156, 160, 162, 164, 169, 287, 294, 306, and 326, respectively); and functionally active fragments and variants thereof.

31. A polypeptide according to Claim 25, wherein said polypeptide is LCR or LCR-like and includes an amino acid sequence shown in Figure 66 hereto (Sequence ID No: 186); and functionally active fragments and variants thereof.

32. A polypeptide according to Claim 25, wherein said polypeptide is F3'5'H or F3'5'H-like and includes an amino acid sequence selected from the group consisting of sequences shown in Figures 69 and 71 hereto (Sequence ID Nos: 195 and 197, respectively); and functionally active fragments and variants thereof.

33. A polypeptide according to Claim 25, wherein said polypeptide is F3H or F3H-like and includes an amino acid sequence selected from the group consisting of sequences shown in Figures 74, 77, 79, 108, 112 and 173 hereto (Sequence ID Nos: 203, 246, 248, 299, 304, and 328, respectively); and functionally active fragments and variants thereof.

34. A polypeptide according to Claim 25, wherein said polypeptide is F3'H or F3'H-like and includes an amino acid sequence selected from the group consisting of sequences shown in Figure 81 hereto (Sequence ID No: 250); and functionally active fragments and variants thereof.

35. A polypeptide according to Claim 25, wherein said polypeptide is PAL or PAL-like and includes an amino acid sequence selected from the group consisting of sequences shown in Figures 84, 87, 90, 92, 94, 96, 178, 183 and 188 hereto (Sequence ID Nos: 254, 259, 269, 271, 273, 275, 330, 332, and 334, respectively); and functionally active fragments and variants thereof.

36. A polypeptide according to Claim 25, wherein said polypeptide is VR or VR-like and includes an amino acid sequence selected from the group consisting of sequences shown in Figures 99 and 193 hereto (Sequence ID Nos: 279 and 336, respectively); and functionally active fragments and variants thereof.

INTERNATIONAL SEARCH REPORT

International application No.
PCT/AU02/01345

A. CLASSIFICATION OF SUBJECT MATTER

Int. Cl. [7]: C12N 15/29, 15/52, 15/53, 15/55, 15/60, 15/61; A01H 5/00

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
SEE ELECTRONIC DATABASE BOX BELOW

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched
SEE ELECTRONIC DATABASE BOX BELOW

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)
PEPTIDE DATABASES (SWISSPROT, GENBANK, EMBL, PIR) DGENE: SEQ ID NOS 2,9,14,18,24,65,70,79,92,96,109,111,118,136,148,154,156,160,162,164,169,186,195,197,203,246,248,250,279,287,294,299,304,308,310,312,314,318,320,322,324,326,328,334

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
PX	WO 02 18604 A (THE SALK INSTITUTE FOR BIOLOGICAL STUDIES) 7 March 2002 Fig 1D shares ~92% identity with SEQ ID NOS 2, 9, 14, 308	1-3, 13-19, 25, 27
X	US 6 054 636 A (FADER GM) 25 April 2000 Fig 2 shares ~81% identity with SEQ ID NOS 2, 9, 14, 308	1-3, 13-27

[X] Further documents are listed in the continuation of Box C   [X] See patent family annex

* Special categories of cited documents:

- "A" document defining the general state of the art which is not considered to be of particular relevance
- "E" earlier application or patent but published on or after the international filing date
- "L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
- "O" document referring to an oral disclosure, use, exhibition or other means
- "P" document published prior to the international filing date but later than the priority date claimed
- "T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
- "X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
- "Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
- "&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
28 November 2002	05 DEC 2002

Name and mailing address of the ISA/AU	Authorized officer
AUSTRALIAN PATENT OFFICE PO BOX 200, WODEN ACT 2606, AUSTRALIA E-mail address: pct@ipaustralia.gov.au Facsimile No. (02) 6285 3929	**TERRY MOORE** Telephone No : (02) 6283 2632

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WO 99 36543 A (PIONEER HI-BRED INTERNATIONAL, INC.) 22 July 1999 SEQ ID NO 2 shares ~81% identity with SEQ ID NOS 2, 9, 14, 308	1-3, 13-17, 21, 22, 24-27
X	Genbank Acc No AAB41524 chalcone isomerase (*Medicago sativa*) 29 January 1997 (See also Medline Abstract 8193301) 92% identity with SEQ ID NOS 18, 310	1-3, 13-18, 24-27
X	Genbank Acc No CAA74847 anther-specific protein (*Nicotiana sylvestris*) (See also Medline Abstract 99084767) 83% identity with SEQ ID NOS 96, 322 and 67% identity with SEQ ID NOS 318, 70	1, 2, 4, 13-18, 24-26, 28
X	Genbank Acc No CAC14061 chalcone synthase (*Ruta graveolens*) 27 October 2000 ~87% identity with SEQ ID NOS 24, 65, 79, 92, 102, 107, 314, 316, 320, 324	1, 2, 4, 13-18, 24-26, 28
X	Genbank Acc No AAB41556 chalcone reductase (*Medicago sativa*) 30 January 1997 ~95% identity with SEQ ID NOS 109, 118, 312	1, 2, 5, 13-18, 20, 24-26, 29
X	Genbank Acc No CAA11226 chalcone reductase (*Sesbania rostrata*) 3 July 2001 90% identity with SEQ ID NO 111	1, 2, 5, 13-18, 24-26, 29
X	Genbank Acc No AAK52955 dihydro-flavonoid reductase-like protein (*Zea mays*) 14 May 2001 69% identity with SEQ ID NOS 287, 160 and 53% identity with SEQ ID NO 148	1, 2, 6, 13-18, 24-26, 30
PX	WO 02 063021 A (PIONEER HI-BRED INTERNATIONAL, INC.) 15 August 2002 SEQ ID NO 1 shares 74% identity with SEQ ID NOS 148, 160, 287	1, 2, 6, 13-18, 24-26, 30
X	Genbank Acc No AAD54273 dihydroflavonol-4-reducatse DFR1 (*Glycine max*) 10 September 1999 81 % identity with SEQ ID NO 169	1, 2, 6, 13-18, 24-26, 30
X	WO 95 27790 A (CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE) 19 October 1995 SEQ ID NO 1 shares 43% identity with SEQ ID NOS 136, 156, 326, 39% identity with SEQ ID NO 154 and 51% identity with SEQ ID NO 294.	1, 2, 6, 13-18, 24-26, 30
PX	WO 02 10210 A (BAYER AKTIENGESELLSCHAFT) 7 February 2002 SEQ ID NO 2329 shares 51% identity with SEQ ID NO 164, SEQ ID NO 636 shares 62% identity with SEQ ID NO 186, SEQ ID NO 1573 shares 49% identity with SEQ ID NOS 246, 248, SEQ ID NO 2091 shares 57% identity with SEQ ID NOS 304, 299	1, 2, 7, 9, 13-17, 24-26, 31, 32
X	WO 99 14351 (E.I. DU PONT DE NEMOURS AND COMPANY) 25 March 1999 SEQ ID NO 1 shares 56% identity with SEQ ID NO 162	1, 2, 6, 13-17, 21, 22, 24-26, 30

INTERNATIONAL SEARCH REPORT	International application No. PCT/AU02/01345

C (Continuation)	DOCUMENTS CONSIDERED TO BE RELEVANT	
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
PX	WO 02 26994 A (AGRICULTURE VICTORIA SERVICES PTY LTD) 4 April 2002 Fig 39 shares 72% identity with SEQ ID NO 294	1, 2, 6, 13-18, 21, 22, 24-26, 30
X	WO 97 12982 A (CENTRE NATIONAL DE LA RECHERCHE SCEINTIFIQUE) 10 April 1997 SEQ ID NO 5 shares 55% identity with SEQ ID NO 294	1, 2, 6, 13-17, 24-26, 30
X	Genbank Acc No CAA80265 flavonoid 3',5'-hydroxylase (*Petunia x hybrida*) 7 December 1993 70% identity with SEQ ID NO 197	1, 2, 8, 13-18, 24-26, 32
X	EP 1 033 405 (CERES INCORPORATED) 6 September 2000 SEQ ID NO 49742 shares 57% identity with SEQ ID NO 195	1, 2, 8, 13-18, 24-26, 32
X	Genbank Acc No AAF23859 DFR-like protein (*Arabidopsis thaliana*) 11 January 2000 61% identity with SEQ ID NO 186	1, 2, 7, 13-18, 24-26, 31
X	Genbank Acc No BAB01697 oxidase-like protein (*Arabidopsis thaliana*) 27 December 2000 50% identity with SEQ ID NOS 246, 248	1, 2, 9, 13-18, 24-26, 33
X	TREMBL Acc No CAB63776 F3'H1 protein (*Glycine max*) 1 May 2000 85% identity with SEQ ID NO 328, 203	1, 2, 9, 13-18, 24-26, 33
X	Genbank Acc No CAB78172 flavanone 3-beta-hydroxylase (*Arabidopsis thaliana*) 16 March 2000 57% identity with SEQ ID NOS 304, 299	1, 2, 9, 13-17, 24-26, 33
X	Genbank Acc No AAG49298 Flavonoid 3'-hydroxylase (*Callistephus chinensis*) 16 January 2001 68% identity with SEQ ID NO 250	1, 2, 10, 13-18, 20, 24-26, 34
X	Genbank Acc No AAA99500 Phenylalanine ammonia lyase (*Stylosanthes humilis*) 15 May 1996 88% identity with SEQ ID NOS 254, 259, 269, 271, 273, 275, 330, 332, 334	1, 2, 11, 13-18, 24-26, 35
X	Genbank Acc No CAA41169 phenylalanine ammonia lyase (*Medicago sativa*) 5 May 1995 87% identity with SEQ ID NOS 254, 259, 269, 271, 273, 275, 330, 332, 334	1, 2, 11, 13-18, 24-26, 35
X	Genbank Acc No AAB41550 vestitone reductase (*Medicago sativa*) 30 January 1997 (See also Medline abstract 7625843) 95% identity with SEQ ID NOS 336, 279	1, 2, 12-18, 21, 22, 24-26, 36

INTERNATIONAL SEARCH REPORT

International application No.
PCT/AU02/01345

Box I **Observations where certain claims were found unsearchable (Continuation of item 2 of first sheet)**

This international search report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos :

   because they relate to subject matter not required to be searched by this Authority, namely:

2. ☐ Claims Nos :

   because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:

3. ☐ Claims Nos :

   because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a)

Box II **Observations where unity of invention is lacking (Continuation of item 3 of first sheet)**

This International Searching Authority found multiple inventions in this international application, as follows:

See supplemental Box

1. ☒ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims
2. ☐ As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.
3. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:
4. ☐ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Remark on Protest ☐ The additional search fees were accompanied by the applicant's protest.

☒ No protest accompanied the payment of additional search fees.

Supplemental Box 1
(To be used when the space in any of Boxes I to VIII is not sufficient)

Continuation of Box No: II (lack of unity)

The international application does not comply with the requirements of unity of invention because it does not relate to one invention or to a group of inventions so linked as to form a single general inventive concept. The fundamental test for unity of invention is specified in Rule 13.2 of the Regulations under the PCT.

"Where a group of inventions is claimed in one and the same international application, the requirement of unity of invention referred to in Rule 13.1 shall be fulfilled only where there is a technical relationship among those inventions involving one or more of the same or corresponding special technical features. The expression "special technical feature" shall mean those technical features that define a contribution which each of the claimed inventions, considered as a whole, make over the prior art."

The problem addressed by the application is the modification of flavonoid biosynthesis (see page 2, line 29-page 3, line 7). The solution provided by the claims resides in the use of 56 specific polypeptides from clover, medic, ryegrass or fescue species (claim 25) and the nucleic acids or fragments coding for these polypeptides (claim 1). These 56 specific polypeptides fall within the following ten groups:

1.	Chalcone isomeras (CHI)	(SEQ IDS 2,9,14,18,308,310)
2.	Chalcone synthase (CHS)	(SEQ IDS 24,65,70,79,92,96,102,107,314,316,318,322,324)
3.	Chalcone reductase (CHR)	(SEQ IDS 109,111,118,312)
4.	Dihydroflavonol 4-reductase (DFR)	(SEQ IDS 136,148,154,156,160,162,164,169,287,294,326)
5.	Leucoanthocyanidin reductase (LCR)	(SEQ ID 186)
6.	Flavonoid 3',5' hydrolase (F3'5'H)	(SEQ IDS 195,197)
7.	Flavanone 3-hydrolase (F3H)	(SEQ IDS 203,246,248,299,304,328)
8.	Flavonoid 3'-hydroxylase (F3'H)	(SEQ ID 250)
9.	Phenyalanine ammonia-olyase (PAL)	(SEQ IDS 254,259,269,271,273,275,330,332,334)
10.	Vestitone reductase (VR)	(SEQ IDS 279,336)

The application acknowledges that representatives of these enzyme species, and the nucleotides that encode them, are known and have been isolated from other plant species (see page 2 lines 29-31). This is supported by the following documents, which disclose the isolation and characterisation of a number of these enzymes from a range of species, including clover and medic and rye.

EMBL CAA63306	*Secale cereale* chalcone synthase (CHS) (5 March 1999)
GENBANK AAA17993	*Trifolium subterraneum* phenylalanine ammonia-lyase (PAL) (10 May 1994)
PIR S66262	*Medicago sativa* vestitone reductase (VS) (12 November 1999)
SWISSPROT P51109	*Medicago sativa* dihydroflavanol 4-reductase (DHR) (1 October 1996)
SWISSPROT P51088	*Trifolium subterraneum* chalcone synthase (CHS) (15 July 1999)
PIR S44371	*Medicago sativa* chalcone isomerase (CHI) (16 July 1999)
Medline Abstract 11164576	altered pigmentation using CHS and DFR
Medline Abstract 7981963	altered plant stress response using CHS and PAL

These documents disclose not only flavonoid biosynthesis enzymes in a range of plant species they also disclose manipulation of these sequences to modify features such as pigmentation and stress responses. Thus features such as the 10 listed enzyme families, the isolation of representative of these families from a range of plant species and modification of flavonoid pathway enzymes are known. Furthermore, the enzymes isolated from the 4 specific species: clover, medic, ryegrass and fescue, do not appear to contribute any advantage or produce any unexpected result in comparison to known members of the same families isolated from other species. Thus these features are known or are obvious and cannot be regarded as "special technical features" conferring unity on the separate inventions.

Supplemental Box 1
(To be used when the space in any of Boxes I to VIII is not sufficient)

Continuation of Box No: II (lack of unity)

In the absence of an obvious special technical feature, it is appropriate to use the Markush approach.

Claims 1 and 26 are written as claims directed to alternatives, in a so-called Markush style of drafting. The application of the test for Markush claims gives the following result:

(A) the common property is modification of the flavonoid biosynthesis pathway.

(B) (1) no common structure is evident as the structures of the polypeptides are not revealed

(B) (2) there is no single recognised class of compounds embracing all the polypeptides, as the polypeptides belong to different classes ie CHI, CHS, CHR, DFR, LCR F3'5'H, F3H, F3'H, PAL,VR, each carrying out different biological functions.

The species of origin of the polypeptides does not provide a legitimate classification as proteins are primarily classified by their activity not their origin. Thus the polypeptides can be grouped into 10 classes CHI, CHS, CHR, DFR, LCR F3'5'H, F3H, F3'H, PAL, VR, based on their activities, and represent 10 different inventions. Each of these inventions can only be searched using independent search strategies and thus each search requires significant additional effort.

As a service to the Applicant, multiple inventions, as specified by the Applicant, were searched for a single search fee with the proviso that the total number of amino acid sequences associated with the combination of inventions was no greater than 10. This offer was independent of unity consideration and was provided solely as a service to the Applicant. For five additional search fees all the inventions, totalling 56 sequences were searched.

This Annex lists the known "A" publication level patent family members relating to the patent documents cited in the above-mentioned international search report. The Australian Patent Office is in no way liable for these particulars which are merely given for the purpose of information.

Patent Document Cited in Search Report		Patent Family Member					
US	6 054 636	AU	94934/98	EP	1 015 614	WO	99 14351
WO	99 36543	AU	22321/99	EP	1 045 909		
WO	99 14351	AU	94934/98				
WO	97 12892	EP	0 853 672	FR	2739395		

END OF ANNEX